

FINDPATTERNS on geneseq: allowing 0 mismatches				May
1	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		1
<i>accion #</i>				
1	AAW32305	ck: 992 len: 542 1 Aaw32305 Arabidopsis thaliana inorganic phc		1
<i>to match 49 to 01 fraction</i>				
1	130: GVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) - pattern matched (L)XFX(R)(F)WXX(F) - pattern matched LCFFRFWLGF GIGGD		1
1	AAW32302	ck: 605 len: 534 1 Aaw32302 Arabidopsis thaliana inorganic phc		1
<i>matching portion of database seq</i>				
1	130: AVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF GIGGD		1
1	AAW32303	ck: 4821 len: 521 1 Aaw32303 Arabidopsis thaliana inorganic phc		1
1	130: GVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF GIGGD		1
1	AAW32304	ck: 8027 len: 524 1 Aaw32304 Arabidopsis thaliana inorganic phc		1
1	130: GVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF GIGGD		1
1	AAW32301	ck: 8746 len: 524 1 Aaw32301 Arabidopsis thaliana inorganic phc		1
1	130: GVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF GIGGD		1
1	AAW30486	ck: 8893 len: 135 1 Aaw30486 Flea saliva protein fspn (Pfspn6-1		1
1	108: HFDAM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(Q) PGFSRFWNPO QCPAY		1
1	AAW36137	ck: 605 len: 534 1 Aaw36137 A. thaliana inorganic phosphate tr		1
1	130: AVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF GIGGD		1
1	AAW82377	ck: 8893 len: 135 1 Aaw82377 Flea saliva protein Pfspn6-135. 4/		1
1	108: HFDAM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(Q) PGFSRFWNPO QCPAY		1
1	AAW82382	ck: 6654 len: 375 1 Aaw82382 Flea saliva protein Pfspn6-375. 4/		1
1	134: HFDAM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(Q) PGFSRFWNPO QCPAY		1
1	AAW82384	ck: 2624 len: 356 1 Aaw82384 Flea saliva protein Pfspn6-356. 4/		1
1	115: HFDAM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(Q) PGFSRFWNPO QCPAY		1
1	AAW82385	ck: 309 len: 355 1 Aaw82385 Flea saliva protein Pfspn6-357.		1
1	114: HFDAM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(Q) PGFSRFWNPO QCPAY		1
1	AAW82322	ck: 4400 len: 19 1 Aaw82322 p53 homologue TIP 12/1 peptide.		1
1	6: PPLSM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PRFMDYWEGL NENG		1
1	AAW82320	ck: 4400 len: 19 1 Aaw82320 p53 homologue TIP 12/1 peptide.		1
1	6: PPLSM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PRFMDYWEGL NENG		1
1	AAW58852	ck: 5772 len: 130 1 Aaw58852 Human O289_1 secreted protein.		1
1	119: VISHL	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (C)XFX(S)(F)WXX(L) CHFSSFWLGL PA		1
1	AAW56025	ck: 9188 len: 316 1 Aaw56025 phosphate starvation-induced pr		1
1	31: SVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF GIGGD		1
1	AAW37220	ck: 5978 len: 12 1 Aaw37220 MDM2 binding peptide unique pha.		1
1	2: M	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PRFMDYWEGL N		1
1	AAW37224	ck: 9428 len: 15 1 Aaw37224 MDM2 binding peptide unique pha		1
1	5: PRPA	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(Y)WXX(L) LVFADYWEGL Y		1
1	AAW37221	ck: 6151 len: 12 1 Aaw37221 MDM2 binding peptide unique pha		1
1	2: V	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(D)(Y)WXX(O) QNFIDYWTQO F		1
1	AAW37222	ck: 5993 len: 12 1 Aaw37222 MDM2 binding peptide unique pha		1
1	3: TG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(H)(Y)WXX(F) PAFTHYWATF		1
1	AAW37223	ck: 9093 len: 15 1 Aaw37223 MDM2 binding peptide unique pha		1
1	5: IDRA	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(H)WXX(L) PTFRDHWFL V		1
1	AAW37225	ck: 8833 len: 15 1 Aaw37225 MDM2 binding peptide unique pha		1

This Page Blank (uspto)

we accession # to watch
citation to sequence

```
11AA_SEQUENCE 1.0
ID AAW32305 standard; Protein; 542 AA.
XX
AC AAW32305;
XX
DT 27-APR-1998 (first entry)
XX
DE Arabidopsis thaliana inorganic phosphate transporter 5.
XX
KW Columbia strain; inorganic phosphate transporter 5; IPT5;
KM accelerated phosphate uptake; tobacco plant.
XX
OS Arabidopsis thaliana.
XX
PN WO9735984-A1.
XX
PD 02-OCT-1997.
XX
PF 24-MAR-1997; 97WO-JP00975.
XX
PR 25-MAR-1996; 96JP-0094790.
XX
PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.
XX
PI Mitsuoka N, Okumura S, Shibata D, Shirano Y;
XX
PS WPI: 1997-489647/45.
XX
DR N-PSDB; AAT91326.
XX
PT DNA encoding a plant phosphate transporter protein - useful for
PT producing e.g. tobacco plants with increased phosphate uptake and
PT accelerated growth
XX
PS Claim 2; Pages 80-83; 96pp; Japanese.
XX
CC The present sequence is Arabidopsis thaliana (Columbia strain)
CC inorganic phosphate transporter 5 (IPT5), useful for accelerating
CC the uptake of phosphate by plants, e.g. tobacco plants.
XX
SQ Sequence 542 AA;

AAW32305 Length: 542 May 30, 2002 09:25 Type: P Check: 992 ..

1 MAKKGKVLN ALDAKTKQWY HFTAIYIAGM GFTDAYDLF SISLVTKLLG
51 RIYHVDSSK KPGTLPNVA AAVNGVAFCG TLAGLFFGW LGDKLGRKKV
101 YGITLMLVNL CSLGSLSPG HSANGVMATL CFFRFLGFG IGDYPLSAT
151 IMSEYANKKT RGAFAIAVFA MGGFGLAGG IVSLIVSSVF DHAFAPTYE
201 VDPVGSTVPQ ADYVWRIYLM FGAIPALITV YWRMKMPETA RYALVARNT
251 KOAASDMKV LQVDLIAEEE AOSNSNSNP NTFGLFTRE FARRHGLHLL
301 GTTTWFLLD IAYYSSNLFQ KDIYTAIGWI PAETMNAIH EVTVSKAQT
351 LIALGTVPG YWFTVAFIDI LGRFFIOLMG FIFMTIFMEA LAIPYDHMRH
401 RENRIGFLIM YSLTMFANF GPNATTFVVP AEIIPARLRS TCHGISAASG
451 KAGAIVGAFG FLYAAOSSDS EKTDAGYPPG IGVNSLML ACVNLGIVF
501 TLLVPESKGR SLEISREDE EQSGDTFVE MTVANSGRKV PV

11AA_SEQUENCE 1.0
ID AAW32302 standard; Protein; 534 AA.
XX
AC AAW32302;
XX
DT 27-APR-1998 (first entry)
XX
DE Arabidopsis thaliana inorganic phosphate transporter 4.
```

too late

```
XX
KM Columbia strain; inorganic phosphate transporter 4; IPT4;
KW accelerated phosphate uptake; tobacco plant.
XX
OS Arabidopsis thaliana.
XX
PN WO9735984-A1.
XX
PD 02-OCT-1997.
XX
PF 24-MAR-1997; 97WO-JP00975.
XX
PR 25-MAR-1996; 96JP-0094790.
XX
PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.
XX
PI Mitsuoka N, Okumura S, Shibata D, Shirano Y;
XX
PS WPI: 1997-489647/45.
XX
DR N-PSDB; AAT91322, AAT91325.
XX
PT DNA encoding a plant phosphate transporter protein - useful for
PT producing e.g. tobacco plants with increased phosphate uptake and
PT accelerated growth
XX
PS Claim 2; Pages 72-74; 96pp; Japanese.
XX
CC The present sequence is Arabidopsis thaliana (Columbia strain)
CC inorganic phosphate transporter 4 (IPT4), useful for accelerating
CC the uptake of phosphate by plants, e.g. tobacco plants.
XX
SQ Sequence 534 AA;

AAW32302 Length: 534 May 30, 2002 09:25 Type: P Check: 605 ..

1 MAREQLQVLN ALDVAKTKQWY HFTAILIAGM GFTDAYDLF CISLVTKLLG
51 RIYHVEGAQ KPGTLPNVA AAVNGVAFCG TLAGLFFGW LGDKLGRKKV
101 YGMTLMVNL CSIASGLSPG HEKAVMATL CFFRFLGFG IGDYPLSAT
151 IMSEYANKKT RGAFAVSAVFA MGGFGLMAGG IFAILISSAF EAKFSPAYA
201 DDALGSTIPQ ADYVWRIYLM AGAIPAMTY YSRKMPETA RYALVAKDA
251 KOAASDMKV LQVEIEPEQQ KLEISKEKS KAFGLESKF MSRHGLHLLG
301 TTSTWFLLDI AFYSONLFQK DIFSAIGWIP PAOSMAIOE VFKIRAQTL
351 IALGSTVPG YWFTVAFIDYI GRFAIOMMGF FEMTVFMFAL AIPYHMTTHK
401 ENRIGFVIM YSLTFANFG PNATTFVVP AEIIPARFST CHGISAASGK
451 LGAMVGAFCF LYLAQNPDKD KTDAGYPPGI GVRNSLIVLG VVNFGLILFT
501 FLVPESKGRS LEEMSGENED NENSNDSTRT VPIV

11AA_SEQUENCE 1.0
ID AAW32303 standard; Protein; 521 AA.
XX
AC AAW32303;
XX
DT 27-APR-1998 (first entry)
XX
DE Arabidopsis thaliana inorganic phosphate transporter 3.
XX
KW Columbia strain; inorganic phosphate transporter 3; IPT3;
KM accelerated phosphate uptake; tobacco plant.
XX
OS Arabidopsis thaliana.
XX
PN WO9735984-A1.
XX
```

This Page Blank (uspto)

1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(R)(F)WXX(L) PAFSRFWSDL	SAGAH	1	2: Q PTFSDYWKL	P
1	AAW37226 ck: 4557 len: 10 1 Aaw37226 MDM2 binding peptide phage consens (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(D)(Y)WXX(L) PXFXYWXXL		1	AAW37190 ck: 6124 len: 12 1 Aaw37190 Human oncogenic protein MDM2 bl (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFX(D)(Y)WXX(L) ETFSDYWKLL	P
1	AAW37195 ck: 5993 len: 12 1 Aaw37195 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(H)(Y)WXX(F) PAFTHYWATE		1	AAW37191 ck: 6151 len: 12 1 Aaw37191 Human oncogenic protein MDM2 bl (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)xFX(D)(Y)WXX(Q) QNFIDYWTQQ	F
1	3: TG AAW37196 ck: 5978 len: 12 1 Aaw37196 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(D)(Y)WXX(L) PRFMDYWEGL		1	AAW37192 ck: 9093 len: 15 1 Aaw37192 Human oncogenic protein MDM2 bl (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(D)(H)WXX(L) PTFRDHWFAL	V
1	2: M AAW37197 ck: 8253 len: 14 1 Aaw37197 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(D)(Y)WXX(L) PTFSDYWKLL	N	1	AAW37193 ck: 9428 len: 15 1 Aaw37193 Human oncogenic protein MDM2 bl (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFX(D)(Y)WXX(L) LVEADYWETL	Y
1	4: CGQ AAW37203 ck: 634 len: 16 1 Aaw37203 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFX(D)(Y)WXX(L) ETFSDYWKLL	P	1	AAW37194 ck: 8833 len: 15 1 Aaw37194 Human oncogenic protein MDM2 bl (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(R)(F)WXX(L) PAFSRFWSDL	SAGAH
1	6: SGSGQ AAW37204 ck: 1571 len: 28 1 Aaw37204 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(D)(Y)WXX(L) PRFMDYWEGL	NKQIK	1	AAW37170 ck: 5978 len: 12 1 Aaw37170 Human oncogenic protein MDM2 bl (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(D)(Y)WXX(L) PRFMDYWEGL	N
1	3: SM AAW37205 ck: 8233 len: 31 1 Aaw37205 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(D)(Y)WXX(L) PRFMDYWEGL	NXK	1	AAW37171 ck: 6146 len: 12 1 Aaw37171 Human oncogenic protein MDM2 bl (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	19: LAPXM AAW37182 ck: 5993 len: 12 1 Aaw37182 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(H)(Y)WXX(F) PAFTHYWATE		1	AAW37172 ck: 9428 len: 15 1 Aaw37172 Human oncogenic protein MDM2 bl (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFX(D)(Y)WXX(L) LVEADYWETL	Y
1	3: TG AAW37183 ck: 5978 len: 12 1 Aaw37183 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(D)(Y)WXX(L) PRFMDYWEGL	N	1	AAW46748 ck: 562 len: 534 1 Aaw46748 Arabidopsis thaliana phosphate (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFX(R)(F)WXX(F) LCFFRFWLG	GIGGD
1	2: M AAW37184 ck: 6146 len: 12 1 Aaw37184 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFX(D)(Y)WXX(L)		1	AAW46749 ck: 4793 len: 538 1 Aaw46749 Tomato phosphate transporter 1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFX(R)(F)WXX(F) LCFFRFWLG	GIGGD
1			1	129: GVMFT (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFX(D)(Y)WXX(L)	

1	AAW46750	ck: 3154	len: 528	1	Aaw46750 Tomato phosphate transporter 1 prd				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(L)XFX(R)(F)Wxx(F)				
1	135: GVMTT				LCFFRFNLGF			GIGGD	
1	AAW37145	ck: 5772	len: 130	1	Aaw37145 Human O289_1 protein. 6/1998				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(C)XFX(S)(F)Wxx(L)				
1	119: VISHL				CHFSFWIGL			PA	
1	AAV30723	ck: 7459	len: 182	1	Aay30723 Amino acid sequence of a human sec				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(L)XFX(D)(H)Wxx(F)				
1	124: LQPTR				LNFKDHWITF			SVASA	
1	AAV31622	ck: 7787	len: 284	1	Aay31622 Human oxidised LDL receptor (HOLR)				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(E)XFX(S)(Y)Wxx(L)				
1	195: SOSYS				EFFYSYWTGL			LRPDS	
1	AAV38483	ck: 9545	len: 86	1	Aay38483 Human secreted protein encoded by				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(L)XFX(D)(F)Wxx(L)				
1	19: MKKGG				LYFSDFWNKL			DVGAI	
1	AAV38484	ck: 8317	len: 27	1	Aay38484 Human secreted protein encoded by				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(L)XFX(D)(F)Wxx(L)				
1	11: MKKGG				LYFSDFWNKL			DVGAI	
1	AAV36885	ck: 7481	len: 945	1	Aay36885 Protein involved in intermediate p				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(E)XFX(S)(F)Wxx(F)				
1	674: RSSKM				ELFLSFWTAF			SLEFLP	
1	AAV38424	ck: 463	len: 509	1	Aay38424 Human secreted protein. 9/1999				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(L)XFX(D)(F)Wxx(L)				
1	103: MKKAA				LYFSDFWNKL			DVGAI	
1	AAV34694	ck: 6409	len: 945	1	Aay34694 C. pneumoniae protein involved in				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(E)XFX(S)(Y)Wxx(F)				
1	675: RSSKM				ELFLSYWSGF			IPHLN	
1	AAV05317	ck: 2849	len: 280	1	Aay05317 Human secreted protein bn97_1. 6/1				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(E)XFX(S)(Y)Wxx(L)				
1	191: SOSYS				EFFYSYWTGL			LRPDS	
1	AAW67815	ck: 5177	len: 57	1	Aaw67815 Human secreted protein encoded by				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				

	35: DGTPA				(Q)XFX(D)(F)Wxx(L)			ISADP	
					QAFWDFWSHL				
1	AAB56913	ck: 7508	len: 275	1	Aab56913 Human prostate cancer antigen p				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(P)XFX(R)(F)Wxx(L)				
1	253: RHLDR				PVFLRFWETL			DRYMV	
1	AAB58349	ck: 8661	len: 100	1	Aab58349 Lung cancer associated polypept				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(L)XFX(C)(H)Wxx(L)				
1	28: IIXTV				LIFECHWDCL			SDTLH	
1	AAB03957	ck: 4914	len: 531	1	Aab03957 Consensus human mesenchymal ste				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(P)XFX(R)(F)Wxx(L)				
1	509: RHLDR				PVFLRFWETL			DRYMV	
1	AAB36865	ck: 5055	len: 1,503	1	Aab36865 Human putative Ca2+ channel pro				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(L)XFX(D)(F)Wxx(L)				
1	862: MKKAA				LYFSDFWNKL			DVGAI	
1	AAB12123	ck: 1439	len: 344	1	Aab12123 Hydrophobic domain protein from				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(L)XFX(H)(F)Wxx(F)				
1	279: FLNPF				LEFGHFQLEF			NALTL	
1	AAB41885	ck: 2227	len: 251	1	Aab41885 Human ORF ORF1649 polypeptide				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(C)XFX(C)(F)Wxx(L)				
1	79: VYVCV				CHFVCFWVCL			SACLC	
1	AAB42619	ck: 2818	len: 280	1	Aab42619 Human ORF ORF2383 polypeptide				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(E)XFX(S)(Y)Wxx(L)				
1	191: SOSYS				EFFYSYWTGL			LRPDS	
1	AAB32818	ck: 7107	len: 113	1	Aab32818 Eucalyptus grandis transcriptio				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(L)XFX(S)(F)Wxx(L)				
1	72: KIELH				LGFSFWSTL			LDVPP	
1	AAB38371	ck: 3048	len: 88	1	Aab38371 Human secreted protein encoded				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(L)XFX(H)(F)Wxx(F)				
1	23: FLNPF				LEFGHFQLEF			NALTL	
1	AAB17077	ck: 6124	len: 12	1	Aab17077 Mdm/hdm antagonist peptide sequ				
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
					(E)XFX(D)(Y)Wxx(L)				
1	2: Q				ETFSDYMKLL			P	

AAB17078	ck: 6146	len: 12	I Aab17078 Mdm/hdm antagonist peptide sequenc
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
2:	Q		P
PTFSDYWKLL			
AAB17079	ck: 5978	len: 12	I Aab17079 Mdm/hdm antagonist peptide sequenc
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
2:	M		N
PRFMDYMEGL			
AAB17080	ck: 6151	len: 12	I Aab17080 Mdm/hdm antagonist peptide sequenc
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
2:	V		F
QNFIDYWTQQ			
AAB17081	ck: 5993	len: 12	I Aab17081 Mdm/hdm antagonist peptide sequenc
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
3:	TG		
PAFTTHWAITE			
AAB17082	ck: 9093	len: 15	I Aab17082 Mdm/hdm antagonist peptide sequenc
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
5:	IDRA		V
PTFRDHWFAL			
AAB17083	ck: 9428	len: 15	I Aab17083 Mdm/hdm antagonist peptide sequenc
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
5:	PRPA		X
LVFADYWETL			
AAB17084	ck: 8833	len: 15	I Aab17084 Mdm/hdm antagonist peptide sequenc
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
1:			SAGAH
PAFSRFWSDL			
AAB17085	ck: 8896	len: 15	I Aab17085 Mdm/hdm antagonist peptide sequenc
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
1:			SAGAH
PAFSRFWSKL			
AAB17086	ck: 4557	len: 10	I Aab17086 Mdm/hdm antagonist peptide sequenc
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
1:			
PXFXYWXXL			
AAB17089	ck: 6124	len: 12	I Aab17089 Mdm/hdm antagonist peptide sequenc
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
2:	Q		P
ETFSDYWKLL			
AAB17090	ck: 6146	len: 12	I Aab17090 Mdm/hdm antagonist peptide sequenc

1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
2:	Q		P
PTFSDYWKLL			
AAY92944	ck: 5055	len: 1,503	I Aay92944 Human TRPC7 protein. 11/2000
862:	MKKA		DVGAI
LYRSDFWNKL			
AAG17202	ck: 7959	len: 584	I Aag17202 Arabidopsis thaliana protein fr
561:	SAWAM		LSVSF
LIFMCFWVLL			
AAG17203	ck: 1255	len: 549	I Aag17203 Arabidopsis thaliana protein fr
526:	SAWAM		LSVSF
LIFMCFWVLL			
AAG17204	ck: 530	len: 478	I Aag17204 Arabidopsis thaliana protein fr
455:	SAWAM		LSVSF
LIFMCFWVLL			
AAG21029	ck: 5419	len: 504	I Aag21029 Arabidopsis thaliana protein fr
24:	LSLSE		FSPSL
QSFLSEFMSWF			
AAG26394	ck: 3190	len: 329	I Aag26394 Arabidopsis thaliana protein fr
49:	FLRKS		KQRNL
PSFRSHWKS			
AAG26395	ck: 4070	len: 307	I Aag26395 Arabidopsis thaliana protein fr
27:	FLRKS		KQRNL
PSFRSHWKS			
AAG29842	ck: 7775	len: 623	I Aag29842 Arabidopsis thaliana protein fr
353:	PLASD		RKQTF
LSFGRYWNYL			
AAG29843	ck: 6852	len: 550	I Aag29843 Arabidopsis thaliana protein fr
280:	PLASD		RKQTF
LSFGRYWNYL			
AAG29844	ck: 4529	len: 461	I Aag29844 Arabidopsis thaliana protein fr
191:	PLASD		PKQTF
TSFGRYWNYL			

1	AAG36897	ck: 2209	len: 333	1	Aag36897	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(D)(F)Wxx(L) LYPSDFWNKL	DVGAI
1	303: INLSS								
1	AAG36898	ck: 464	len: 318	1	Aag36898	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(S)(Y)Wxx(L) EFFYSYWTGL	LRPDS
1	288: INLSS								
1	AAG36899	ck: 6630	len: 284	1	Aag36899	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(H)(F)Wxx(F) LIFYHEWRYE	HCPAD
1	254: INLSS								
1	AAG37272	ck: 7852	len: 584	1	Aag37272	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGf	GIGGD
1	561: SAWAM								
1	AAG37273	ck: 1025	len: 549	1	Aag37273	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGf	GIGGD
1	526: SAWAM								
1	AAG37274	ck: 969	len: 478	1	Aag37274	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGf	GIGGD
1	455: SAWAM								
1	AAG38413	ck: 3958	len: 261	1	Aag38413	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGf	GIGGD
1	49: FLRKS								
1	AAG38839	ck: 7690	len: 986	1	Aag38839	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGf	GIGGD
1	963: SAWAM								
1	AAG38840	ck: 7810	len: 886	1	Aag38840	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGf	GIGGD
1	863: SAWAM								
1	AAG38841	ck: 4753	len: 859	1	Aag38841	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(D)(H)Wxx(Q) EHFSDHWHKQ	NQDPF
1	836: SAWAM								
1	AAV95439	ck: 5055	len: 1,503	1	Aay95439	Human calcium channel polypeptide	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(E)(F)Wxx(L)	

183: YPTAE	EDFDEFWNVL	RIVGF	1	ABB34423	ck: 1841	len: 113	! Abb34423 Peptide #1929 encoded by human
AAU45398	ck: 183	len: 89	! Aau45398 Propionibacterium acnes immunogeni	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
				(P)xFx(H)(F)Wxx(Q)			
70: QRVSP	PAFPRHMDTL	HEIRA		70: LCEIL			NPTVL
ABG03782	ck: 3344	len: 146	! Abg03782 Novel human diagnostic protein #37	ABBI9835	ck: 1841	len: 113	! Abbi9835 Protein #1834 encoded by probe
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
7: KKIDR	(P)xFx(D)(F)Wxx(F)			70: LCEIL			NPTVL
	PLFKDFWQRF	LDSLK					
ABG03783	ck: 8253	len: 244	! Abg03783 Novel human diagnostic protein #37	AAB86968	ck: 4326	len: 506	! Aab86968 D. melanogaster peptide recepto
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
103: KKIDR	(P)xFx(D)(F)Wxx(F)			323: YPTAE			RIVGF
	PLFKDFWQRF	LDSVK					
ABG05207	ck: 5356	len: 636	! Abg05207 Novel human diagnostic protein #51	AAE11926	ck: 2953	len: 288	! Aae11926 Human CG27 (or C869) receptor p
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
487: SPFER	(Q)xFx(C)(Y)Wxx(L)	ETEHL		199: SQSYS			LRPDS
	QLFVCYWALL						
ABG05563	ck: 5356	len: 636	! Abg05563 Novel human diagnostic protein #55	AAE11932	ck: 2849	len: 280	! Aae11932 Human CG27 (or C869) receptor p
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
487: SPFER	(Q)xFx(C)(Y)Wxx(L)	ETEHL		191: SQSYS			LRPDS
	QLFVCYWALL						
ABG06670	ck: 936	len: 236	! Abg06670 Novel human diagnostic protein #66	AAE11933	ck: 6095	len: 314	! Aae11933 Human CG27 (or C869) receptor p
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
38: GVLWY	(L)xFx(S)(F)Wxx(L)	RSPVP		218: SQSYS			LRPDS
	LTFSSFWAPL						
ABG06767	ck: 5356	len: 636	! Abg06767 Novel human diagnostic protein #67	AAE11934	ck: 89	len: 247	! Aae11934 Human CG27 (or C869) receptor p
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
487: SPFER	(Q)xFx(C)(Y)Wxx(L)	ETEHL		158: SQSYS			LRPDS
	QLFVCYWALL						
ABG17402	ck: 5356	len: 636	! Abg17402 Novel human diagnostic protein #17	AAE11938	ck: 2849	len: 280	! Aae11938 Human lipid metabolism related
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
487: SPFER	(Q)xFx(C)(Y)Wxx(L)	ETEHL		191: SQSYS			LRPDS
	QLFVCYWALL						
ABG29764	ck: 9364	len: 198	! Abg29764 Novel human diagnostic protein #29	AAE11939	ck: 2921	len: 307	! Aae11939 Human lipid metabolism related
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
59: GVTIRP	(E)xFx(H)(F)Wxx(F)	RTKGD		218: SQSYS			LRPDS
	EDFVHFWMRF						
ABB29544	ck: 8372	len: 27	! Abb29544 Peptide #2195 encoded by breast ce	AAE11943	ck: 7726	len: 281	! Aae11943 Human CG27 (or C869) receptor p
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
8: KKIDR	(P)xFx(D)(F)Wxx(F)	LDSLK		185: SQSYS			LRPDS
	PLFKDFWQRF						

42: GSLET	(L)xFx(S)(H)Wxx(Q) LLEFISHWPEQ	SEMAI
AAU27511	ck: 6166 len: 189 i Aau27511 Human G-Protein Coupled Receptor (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(S)(Y)Wxx(Q) LCFMSYWEPPQ	1
167: SYVRV	FFPGG	
AAU29324	ck: 2849 len: 280 i Aau29324 Human PRO polypeptide sequence #30 (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(S)(Y)Wxx(L) EFFYSYWTGL	1
191: SQSYS	LRPDS	
AAU32368	ck: 4809 len: 273 i Aau32368 Novel human secreted protein #2859 (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(D)(F)Wxx(F) PLFKDFWQRF	1
88: KKIDR	LDSVK	
AAM55526	ck: 8372 len: 27 i Aam55526 Human brain expressed single exon (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(D)(F)Wxx(F) PLFKDFWQOF	1
8: KKIDR	LDSLK	
AAM67910	ck: 8372 len: 27 i Aam67910 Human bone marrow expressed probe (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(D)(F)Wxx(F) PLFKDFWQOF	1
8: KKIDR	LDSLK	
AAW79021	ck: 5679 len: 377 i Aam79021 Human protein SEQ ID NO 1683. 11/2 (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(H)(F)Wxx(F) LFFGFHWQLF	1
279: FLLPF	NALTL	
AAM80005	ck: 7978 len: 406 i Aam80005 Human protein SEQ ID NO 3651. 11/2 (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(H)(F)Wxx(F) LFFGFHWQLF	1
342: FLLPF	NALTL	
AAM84808	ck: 1969 len: 36 i Aam84808 Human immune/haematopoietic antigen (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (C)xFx(S)(F)Wxx(F) CCFSSFWKLF	1
15: GLIFC	RTGHL	
AAM93544	ck: 2849 len: 280 i Aam93544 Human polypeptide, SEQ ID NO: 3297 (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(S)(Y)Wxx(L) EFFYSYWTGL	1
191: SQSYS	LRPDS	
AAO01590	ck: 9232 len: 113 i Aao01590 Human polypeptide SEQ ID NO 15482. (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(C)(F)Wxx(L) PGFFCFWAPL	1
59: FLNPP	KKFPF	
AAB47415	ck: 6481 len: 585 i Aab47415 Plant phosphate transporter, Le (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGf	1
130: GVIGT	GIGGD	
AAM15727	ck: 8372 len: 27 i Aam15727 Peptide #2161 encoded by probe (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(D)(F)Wxx(F) PLFKDFWQOF	1
8: KKIDR	LDSLK	
AAM28232	ck: 8372 len: 27 i Aam28232 Peptide #2269 encoded by probe (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(D)(F)Wxx(F) PLFKDFWQOF	1
8: KKIDR	LDSLK	
AAM39779	ck: 8327 len: 80 i Aam39779 Human polypeptide SEQ ID NO 292 (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(H)(F)Wxx(F) LLEFYHWRYP	1
17: GLNVP	HCPAD	
AAM39876	ck: 4914 len: 531 i Aam39876 Human polypeptide SEQ ID NO 302 (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(R)(F)Wxx(L) PVFLRFWETL	1
509: RHLDR	DRYV	
AAM40139	ck: 2511 len: 258 i Aam40139 Human polypeptide SEQ ID NO 328 (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(D)(F)Wxx(F) PLFKDFWQRF	1
73: KKIDR	LDSLK	
AAM41565	ck: 4766 len: 104 i Aam41565 Human polypeptide SEQ ID NO 649 (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(H)(F)Wxx(F) LLEFYHWRYP	1
41: GLNVP	HCPAD	
AAM41925	ck: 2688 len: 207 i Aam41925 Human polypeptide SEQ ID NO 685 (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(D)(F)Wxx(F) PLFKDFWQOF	1
22: KKIDR	LDSLK	
AAU12400	ck: 2849 len: 280 i Aau12400 Human PRO1131 polypeptide sequence (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(S)(Y)Wxx(L) EFFYSYWTGL	1
191: SQSYS	LRPDS	
AAG92957	ck: 4387 len: 592 i Aag92957 C glutamicum protein fragment S (P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(D)(Y)Wxx(L) EERIDYWAL	1
559: GHNVH	LDRV	
AAG92978	ck: 2649 len: 507 i Aag92978 C glutamicum protein fragment S	

1	348: EAAEG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(F)WXX(L) ENFARFWSWL	KSRRD
1	AAG74295	ck: 942 len: 66 1 Aag74295 Human colon cancer antigen protein (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(H)WXX(Q) LFFMHMGAAQ	SWXNL
1	40: QCMHE		
1	AAG89196	ck: 4447 len: 160 1 Aag89196 Human secreted protein, SEQ ID NO:1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LLEFYHFWRYF	HCPAD
1	97: GLNVP		
1	AAU03216	ck: 0 len: 483 1 Aau03216 Fruit fly G protein coupled recept (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(E)(F)WXX(L) EDEDEFWNVL	RILPK
1	323: YPTAE		
1	AAB93312	ck: 4753 len: 732 1 Aab93312 Human protein sequence SEQ ID NO:1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQRF	LDSLK
1	547: KKIDR		
1	AAB94345	ck: 6460 len: 561 1 Aab94345 Human protein sequence SEQ ID NO:1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLEKDFWQRF	LDSLK
1	376: KRIDR		
1	AAB60464	ck: 4282 len: 160 1 Aab60464 Human cell cycle and proliferation (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LLEFYHFWRYF	HCPAD
1	97: GLNVP		
1	AAB60465	ck: 4914 len: 531 1 Aab60465 Human cell cycle and proliferation (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PVFLRFWETL	DRYMY
1	509: RHIDR		
1	AAB65251	ck: 2849 len: 280 1 Aab65251 Human PRO1131 (UNQ569) protein seq (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFYSYWTGL	LRPDS
1	191: SQSYS		
1	AAB50959	ck: 2849 len: 280 1 Aab50959 Human PRO1131 protein. 3/2001 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFYSYWTGL	LRPDS
1	191: SQSYS		
1	AAU19776	ck: 52 len: 394 1 Aau19776 Human novel extracellular matrix p (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LEFGHFWQLF	NALTL
1	304: FLLEP		

1	ABB48814	ck: 2157 len: 510 1 Abb48814 Listeria monocytogenes protein (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(Y)WXX(F) PNFDRYWADF	PHGEL
1	31: AQANK		
1	ABB49398	ck: 1607 len: 289 1 Abb49398 Listeria monocytogenes protein (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(F)WXX(L) EDFLSFWNNL	RMTVT
1	175: HLNFS		
Databases searched:			
EMBL, Release 7.0, Released on 28Apr2002, Formatted on 2May2002			
Total finds: 158			
Total length: 114,001,827			
Total sequences: 766,495			
CPU time: 14:23.28			

11AA_SEQUENCE 1.0

ID AAW32305 standard; Protein; 542 AA.

AC AAW32305;

DT 27-APR-1998 (first entry)

DE Arabidopsis thaliana inorganic phosphate transporter 5.

KW Columbia strain; inorganic phosphate transporter 5; IPT5;

KW accelerated phosphate uptake; tobacco plant.

OS Arabidopsis thaliana.

PN WO9735984-A1.

PD 02-OCT-1997.

PE 24-MAR-1997; 97WO-JP00975.

PR 25-MAR-1996; 96JP-0094790.

PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.

PI Mitsuoka N, Okumura S, Shibata D, Shirano Y;

DR WPI; 1997-489647/45.

DR N-PSDB; AAT91326.

PT DNA encoding a plant phosphate transporter protein - useful for producing e.g. tobacco plants with increased phosphate uptake and accelerated growth

PS Claim 2; Pages 80-83; 96pp; Japanese.

CC The present sequence is Arabidopsis thaliana (Columbia strain)

CC inorganic phosphate transporter 5 (IPT5), useful for accelerating the uptake of phosphate by plants, e.g. tobacco plants.

SQ Sequence 542 AA;

AAW32305 Length: 542 May 30, 2002 09:25 Type: P Check: 992 ..

1 MAKKGKVLN ALDAKTQWY HFTAIYIAGM GFTDAYDLF SISLVTKLIG

51 RIYVHVDSK KPGTLPPNVA AAVNGVAFCG TLAGLFFGW LGDKLGRKKV

101 YGITLMLVL CSLGSLFSG HSANGVMATL CFFRWLFGF IGDYPLSAT

151 IMSEYANKKT RGAFAIAVFA MGGFGLAGG IVSLIVSTF DHAFAPIVE

201 VDPVGSTVPQ ADYVWRIVLM FGAIPALTY YWRMKMPETA RYTALVARNT

251 KOAASDMSKV LQVDLIAEEE AQSNSNSNP NTFGLFTRE FARHGLHL

301 GTTTFWFLLD IAYYSSNLFO KDIYTAIGWI PAETMNAIH EVTVSKAQT

351 LIALCGTVPQ YWETVAFIDI LGRFIIQIMG FIFMTIEMFA LAIPYDHMRH

401 RENRIGFLIM YSLTMFFANF GPNATTFVVP AEIFPARLRS TCHGISASG

451 KAGAIVGAFG FLVAAOSSDS EKTDAGYPPG IGVNSLML ACVNFGLIVF

501 TLLVPESKSK SLEISREDE EQSGDITYVE MTVANSGRKV PV

11AA_SEQUENCE 1.0

ID AAW32302 standard; Protein; 534 AA.

AC AAW32302;

DT 27-APR-1998 (first entry)

DE Arabidopsis thaliana inorganic phosphate transporter 4.

XX

KW Columbia strain; inorganic phosphate transporter 4; IPT4;

KW accelerated phosphate uptake; tobacco plant.

OS Arabidopsis thaliana.

PN WO9735984-A1.

PD 02-OCT-1997.

PE 24-MAR-1997; 97WO-JP00975.

PR 25-MAR-1996; 96JP-0094790.

PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.

PI Mitsuoka N, Okumura S, Shibata D, Shirano Y;

DR WPI; 1997-489647/45.

DR N-PSDB; AAT91322, AAT91325.

PT DNA encoding a plant phosphate transporter protein - useful for producing e.g. tobacco plants with increased phosphate uptake and accelerated growth

PS Claim 2; Pages 72-74; 96pp; Japanese.

CC The present sequence is Arabidopsis thaliana (Columbia strain)

CC inorganic phosphate transporter 4 (IPT4), useful for accelerating the uptake of phosphate by plants, e.g. tobacco plants.

SQ Sequence 534 AA;

AAW32302 Length: 534 May 30, 2002 09:25 Type: P Check: 605 ..

1 MARBQLQVLN ALDVAKTQWY HFTAIYIAGM GFTDAYDLF CISLVTKLIG

51 RIYVHVEGAQ KPGTLPPNVA AAVNGVAFCG TLAGLFFGW LGDKLGRKKV

101 YGMTLMVMVL CSIASGLSFG HEPKAVMATL CFFRWLFGF IGDYPLSAT

151 IMSEYANKKT RGAFAVSAVFA MGGFCIMAGG IFALIISSAF EAKFPSPAYA

201 DDALGSTIPQ ADLVWRIILM AGAIPAAMTY YSRSKMPETA RYTALVAKDA

251 KOAASDMSKV LQVEIEPEQK KLEISKEKS KAFGLFSKEF MSRHGLHLIG

301 TTSWTFLLDI AFYSONLFOK DIFSAIGWIP PAOSMNAIQE VFRIARAQTL

351 IALCSTVPQY WFTVAFIDVI GRFAIQMGF FFMVFMFAL AIPYNHWTHK

401 ENRIGFVIMY SLTFFEFANFG PNATTFVPA EITPARFRST CHGISASGK

451 LGAMVGAFCF LYLAQNPKDK KTDAGYPPGI GVRNSLIVLG VVNFLGILFT

501 FLVPEKSKGS LEMSGENED NENSNDSTRT VPIV

11AA_SEQUENCE 1.0

ID AAW32303 standard; Protein; 521 AA.

AC AAW32303;

DT 27-APR-1998 (first entry)

DE Arabidopsis thaliana inorganic phosphate transporter 3.

KW Columbia strain; inorganic phosphate transporter 3; IPT3;

KW accelerated phosphate uptake; tobacco plant.

OS Arabidopsis thaliana.

PN WO9735984-A1.

PD 02-OCT-1997.
XX
XX
PF 24-MAR-1997; 97WO-JP00975.
XX
PR 25-MAR-1996; 96JP-0094790.
XX
XX
PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.
XX
PI Mitsukawa N, Okumura S, Shibata D, Shirano Y;
XX
XX
DR WPI: 1997-489647/45.
DR N-PSDB; AAT91324.
XX
XX
PT DNA encoding a plant phosphate transporter protein - useful for
PT producing e.g. tobacco plants with increased phosphate uptake and
PT accelerated growth
XX
XX
PS Claim 2; Pages 75-77; 96pp; Japanese.
XX
XX
CC The present sequence is Arabidopsis thaliana (columbia strain)
CC Inorganic phosphate transporter 3 (IPT3), useful for accelerating
CC the uptake of phosphate by plants, e.g. tobacco plants.
XX
XX
SQ Sequence 521 AA;

AAW32303 Length: 521 May 30, 2002 09:25 Type: P Check: 4821 ..

1 MADQQLGVLK ALDVAKTQLY HFTAIVIAGM GFTDAYDLF CVSLVTKLLG
51 RLYYFNPTSA KPGSLPPHYA AAVNGVALCG TLAGQLFEGW LGDKLGRKKV
101 YGTLIMMIL CSVASGLSLG NSAKGVMTTL CFFRFLGFG IGDYPLSAT
151 IMSEYANKKT RGAFIAAVEA MGVGILAGG FVALAVSSIF DKKFSPSTYE
201 QDRFLSTPPQ ADYIWRITVM FGALPALTY YWRMKMPETA RYTALVAKNI
251 KOATADMSKV LQTDLEER VEDVDKDPK NYGLFSKEFL RRHGLHLGT
301 TSTWFLDIA FYSQNLFOKD IFSALGWIPK AATMAIHEV FKIRAOQLI
351 ALCSTVPGYW FTVAFIDTIG RFAIQLMGEF MMTVFMEAIA FPNHWILLPD
401 NRIGFVVMYS LTFEFANFGP NATTFIVPAE IFPARLRSTC HGISAATGKA
451 GAIVGAFGFL YAAQPODKTK TDAGYPPGIG VKNSLIMLGV INFVGMLETF
501 LVPEPKGKSL EELSGEAEVD K

!!AA_SEQUENCE 1.0

ID AAW32304 standard; Protein; 524 AA.

AC AAW32304;

DT 27-APR-1998 (first entry)

DE Arabidopsis thaliana Inorganic phosphate transporter 2.

KW Columbia strain; Inorganic phosphate transporter 2; IPT2;
KW accelerated phosphate uptake; tobacco plant.

OS Arabidopsis thaliana.

PN WO9735984-A1.

PD 02-OCT-1997.

PF 24-MAR-1997; 97WO-JP00975.

PR 25-MAR-1996; 96JP-0094790.

PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.

PI Mitsukawa N, Okumura S, Shibata D, Shirano Y;
XX
XX
DR WPI: 1997-489647/45.
DR N-PSDB; AAT91324.
XX
XX
PT DNA encoding a plant phosphate transporter protein - useful for
PT producing e.g. tobacco plants with increased phosphate uptake and
PT accelerated growth
XX
XX
PS Claim 2; Pages 77-80; 96pp; Japanese.
XX
XX
CC The present sequence is Arabidopsis thaliana (columbia strain)
CC Inorganic phosphate transporter 2 (IPT2), useful for accelerating
CC the uptake of phosphate by plants, e.g. tobacco plants.
XX
XX
SQ Sequence 524 AA;

AAW32304 Length: 524 May 30, 2002 09:25 Type: P Check: 8027 ..

1 MAEQQLGVLK ALDVAKTQLY HFTAIVIAGM GFTDAYDLF CVSLVTKLLG
51 RLYYFNPTSA KPGSLPPHYA AAVNGVALCG TLAGQLFEGW LGDKLGRKKV
101 YGTLIMMIL CSVASGLSLG NSAKGVMTTL CFFRFLGFG IGDYPLSAT
151 IMSEYANKKT RGAFIAAVEA MGVGILAGG FVALAVSSIF DKKFAPPIYA
201 VNRLSTPPQ VDIYWRITVM FGALPALTY YWRMKMPETA RYTALVAKNI
251 KOATADMSKV LQTDLEER VEDVDKDPK NYGLFSKEFL RRHGLHLGT
301 TSTWFLDIA FYSQNLFOKD IFSALGWIPK AATMAIHEV FKIRAOQLI
351 ALCSTVPGYW FTVAFIDTIG RFAIQLMGEF MMTVFMEAIA FPNHWIKPE
401 NRIGFVVMYS LTFEFANFGP NATTFIVPAE IFPARLRSTC HGISAAGKA
451 GAIIGAFGFL YAAQPODKTK TDAGYPPGIG VKNSLIMLGV LNFVGMLETF
501 LVPEPKGKSL EELSGEAEVS HDEK

!!AA_SEQUENCE 1.0

ID AAW32301 standard; Protein; 524 AA.

AC AAW32301;

DT 27-APR-1998 (first entry)

DE Arabidopsis thaliana Inorganic phosphate transporter 1.

KW Columbia strain; Inorganic phosphate transporter 1; IPT1;
KW accelerated phosphate uptake; tobacco plant.

OS Arabidopsis thaliana.

PN WO9735984-A1.

PD 02-OCT-1997.

PF 24-MAR-1997; 97WO-JP00975.

PR 25-MAR-1996; 96JP-0094790.

PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.

PI Mitsukawa N, Okumura S, Shibata D, Shirano Y;

DR WPI: 1997-489647/45.

DR N-PSDB; AAT91321, AAT91323.

PT DNA encoding a plant phosphate transporter protein - useful for
PT producing e.g. tobacco plants with increased phosphate uptake and
PT accelerated growth

XX Claim 2; Pages 69-71; 96pp; Japanese.
PS
XX
CC The present sequence is Arabidopsis thaliana (columbia strain)
CC Inorganic phosphate transporter 1 (IPT1), useful for accelerating
CC the uptake of phosphate by plants, e.g. tobacco plants.
XX
SQ Sequence 524 AA;
AAW32301 Length: 524 May 30, 2002 09:25 Type: P Check: 8746 ..
1 MAEQQLGVLK ALDVAKTOLY HFTAVIAGM GFFTDAYDLF CVSLVTKLLG
51 RIYYNPESA KPGSLPHVA AAVNGVALCG TISGQLEFGW LGDKLGRKKV
101 YGLTLVMIL CSVASGLSFG HEAKGVMTTL CFFRWLGFG IGDYPLSAT
151 IMSEYANKKT RGAFAIAVFA MGVGILAGG FVALAVSSIF DKFPAPTYA
201 VNRAALSTPQ VDIWRIIVM FGALPALTY YWRMKMPETA RYTALVAKNI
251 KQATADMSKV LQTDIELEER VEDDVKDPKQ NYGLFSKEFL RHGHLHLGT
301 TSTWFLDLIA FYSQNLFOKD IFSALGWIPK AATMNATHEV FRIARAQTLI
351 ALCSTVPGYW FTVAFIDTIG REKIQNGFF MMTVEMFAIA FPNYHWIKPE
401 NRIGFVVMYS LTFEEFANFGP NATFEIVPAE IFPARLRSTC HGISAAGKA
451 GAIVGAFGFL YAAQSODKAK VDAGYPPGIG VKNSLIMLGV LNFIGMLFTF
501 LVPEPKGKSL EELSGEAEVS HDEK
11AA_SEQUENCE 1.0
ID AAW30486 standard; Protein; 135 AA.
XX
AC AAW30486;
XX
DT 14-APR-1998 (first entry)
XX
DE Flea saliva protein fspn (PfSPN6-135).
XX
KW Flea saliva protein; fspn; allergic dermatitis; allergy;
KM therapy; diagnosis; antibody; PfSPN6-135.
XX
OS Ctenocephalides felis.
XX
PN WO9737676-A1.
XX
PD 16-OCT-1997.
XX
PF 10-APR-1997; 97WO-US05959.
XX
PR 10-APR-1996; 96US-0630822.
XX
PA (HESK-) HESKA CORP.
XX
PI Hunter SW, Slim G, Weber ER;
XX
DR WPI: 1997-512409/47.
DR N-PSDB; AAT92826.
XX
PT New flea saliva proteins - useful for treating allergic dermatitis
PT and as diagnostic reagents
XX
PS Claim 2; Page 156; 179pp; English.
XX
CC This polypeptide comprises a non-full-length flea saliva protein
CC (FSP), denoted PfSPN6-135, that can be used to treat allergic
CC dermatitis. Its amino acid sequence was deduced from nucleic
CC acid nFSPN6-406 (see AAT92826). Claimed FSP polypeptides (see also
CC AAW30480, AAW30483, AAW30487 and AAW30488-91) can be expressed in host
CC cells. The proteins, or their fragments or mimetopes, are used in

CC claimed methods for treating allergic dermatitis in animals, to
CC determine if an animal is susceptible to, or has, allergic
CC dermatitis, and to desensitise a host animal to allergic
CC dermatitis, as well as to monitor progress or effects of treatment.
CC Also contemplated is the in vivo expression of FSPs. FSPs can also
CC be used to raise antibodies useful as immunoassay reagents and for
CC passive immunisation.
XX
SQ Sequence 135 AA;
AAW30486 Length: 135 May 30, 2002 09:25 Type: P Check: 8893 ..
1 MVKGPHEAC NYAGGPOLTT LQEKDSVLTE DGTAYEELG KLLDKVYKKQ
51 LKVDKWDATK TYMAVSTKAM RTKEAALIVG AGLENNPAKA KGNWTOOQLD
101 STHFDAMPGF SRFWNPQCCP AYFRAISLON OKIKK
11AA_SEQUENCE 1.0
ID AAW36137 standard; Protein; 534 AA.
XX
AC AAW36137;
XX
DT 25-MAR-1998 (first entry)
XX
DE A. thaliana inorganic phosphate transporter protein IPT4.
XX
XX Promoter; inorganic phosphate transporter gene; IPT4; expression;
KW regulation; transgenic plant; expression vector.
XX
OS Arabidopsis thaliana.
XX
PN JP09252782-A.
XX
PD 30-SEP-1997.
XX
PF 25-MAR-1996; 96JP-0094856.
XX
PR 25-MAR-1996; 96JP-0094856.
XX
PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
XX
DR WPI: 1997-530154/49.
DR N-PSDB; AAV01455.
XX
PT A promoter for a phosphate transporter gene - useful for regulating
PT heterologous gene expression in transgenic plants
XX
PS Example 2; Page 6-9; 12pp; Japanese.
XX
CC This is the amino acid sequence of the Arabidopsis thaliana inorganic
CC phosphate transporter gene IPT4. The promoter region (bases 1-1875; see
CC also AAV01454) of the encoding gene can be used to regulate the level of
CC expression of a heterologous gene in a plant by introducing, into a
CC plant, a vector having the heterologous gene linked downstream region
CC from the promoter and then regulating extracellular phosphate levels.
XX
SQ Sequence 534 AA;
AAW36137 Length: 534 May 30, 2002 09:25 Type: P Check: 605 ..
1 MAREQLQVLN ALDVAKTQWY HFTAILIAGM GFFTDAYDLF CISTVTKLLG
51 RIYHVEGAQ KPPTLPPNVA AAVNGVAFCG TLAGQLEFGW LGDKLGRKKV
101 YGMTLMWVL CSIASGLSFG HEKAVMATL CFFRWLGFG IGDYPLSAT
151 IMSEYANKKT RGAFSAVFA MGVGIMAGG IFALTISSAF EAKPPSPAYA
201 DDALGSTIPQ ADLWRIILM AGAIPAMTY YSRSKMPETA RYTALVAKDA
251 KQASDMSKV LQVEIEPEQO KLEETSEKS KAFGLFSKEF MSRHGLHLLG

301 TTSTWFLLDI AFYSQNLFOK DIFSAIGWIP PAQSMNAIOE VFKIARAOTL
XX
351 IALCSTVPGY WFTVAFIDVI GRFAIQMGF FEMTVEMFAL AIPYNHWHK
XX
401 ENRIGFVIMY SLTEFFANFG PNATTFVYPA EIEPARFRST CHGISASGK
XX
451 LGAMVGAFGF LYLAQNPPDKD KTDAGYPPGI GVRNSLIYLG VVNFGLILEFT
XX
501 FLVPESEKGRS LEEMSGENED NENSNDSDRT VPIV
XX
11AA_SEQUENCE 1.0
ID AAW82377 standard; Protein; 135 AA.
XX
AC AAW82377;
XX
DT 26-APR-1999 (first entry)
XX
DE Flea saliva protein Pfspn6-135.
XX
KW Flea saliva protein; Pfspn6-135; allergic dermatitis; allergen;
XX
KM allergy; therapy; diagnosis; vaccine; ectoparasite.
XX
OS Ctenocephalides sp.
XX
PN WO9845408-A2.
XX
PD 15-OCT-1998.
XX
PF 15-OCT-1997; 97WO-US18669.
XX
PR 10-APR-1997; 97WO-US05959.
XX
PA (HESK-) HESKA CORP.
XX
PI Weber ER;
XX
DR WPI; 1998-594480/50.
XX
DR N-PSDB; AAV73417.
XX
XX
PT New isolated ectoparasite saliva genes - used to develop products
PT for the diagnosis, prevention, treatment and determining
PT susceptibility to allergic dermatitis
XX
PS Example 10; Page 125; 172pp; English.
XX
XX This is flea saliva protein Pfspn6-135, encoded by a cDNA clone,
CC nfspn6-405 (see AAV73417), that was isolated from a flea salivary
CC gland cDNA library. Pfspn6-135 shows no significant homology to
CC known amino acid sequences. The invention relates to novel methods
CC for isolating ectoparasite saliva proteins (ESPs), including flea
CC saliva proteins. It provides ESps (claimed, see AAW82382-93), nucleic
CC acids encoding them, methods for their recombinant production,
CC therapeutic compositions for treating allergic dermatitis that
CC comprise at least one ESP, as well as assay kits for testing if an
CC animal has, or is susceptible to, allergic dermatitis, and a method
CC of desensitising a host animal to allergic dermatitis using ESps.
CC The ESps can also be used for the production of antibodies useful
CC in diagnosis or in vaccines for passive immunisation against
CC allergic dermatitis.
XX
SQ Sequence 135 AA;
AAW82377 Length: 135 May 30, 2002 09:25 Type: P Check: 8893 ..
1 MVKGPDEAC NYAGPQLTT LOEKDSVLTG DGTETAYELG KLDKRYKKQ
51 LKVDKWDATK TYWAVSTKAM RTKEAALIVG AGLENNPAKA KGNWTQOQLD
101 STHFDAMPGF SRFWNPOQCP AYFRALSLQN QKIKK
11AA_SEQUENCE 1.0
ID AAW82382 standard; Protein; 375 AA.
XX

AC AAW82382;
XX
DT 26-APR-1999 (first entry)
XX
DE Flea saliva protein Pfspn6-375.
XX
KW Flea saliva protein; Pfspn6-375; allergic dermatitis; allergen;
XX
KM allergy; therapy; diagnosis; vaccine; ectoparasite.
XX
OS Ctenocephalides sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Protein /note= "signal peptide"
FT /note= "mature protein"
XX
PN WO9845408-A2.
XX
PD 15-OCT-1998.
XX
PF 15-OCT-1997; 97WO-US18669.
XX
PR 10-APR-1997; 97WO-US05959.
XX
PA (HESK-) HESKA CORP.
XX
PI Weber ER;
XX
DR WPI; 1998-594480/50.
DR N-PSDB; AAV73430 AND AAV73432.
XX
XX
PT New isolated ectoparasite saliva genes - used to develop products
PT for the diagnosis, prevention, treatment and determining
PT susceptibility to allergic dermatitis
XX
PS Claim 2; Page 133-134; 172pp; English.
XX
XX This polypeptide comprises a new full-length flea saliva protein,
CC termed Pfspn6-375. Its amino acid sequence was deduced from a
CC cDNA clone (see AAV73430) isolated from a whole flea cDNA library,
CC and shows no significant homology to known amino acid sequences.
CC The mature fspn6-356 protein (see AAW82384) has been expressed
CC in recombinant Escherichia coli BL-21 cells. The invention is
CC directed to methods for isolating ectoparasite saliva proteins
CC (ESps). It provides ESps (claimed, see AAW82382-93), nucleic acid
CC molecules encoding them, methods for their recombinant production,
CC therapeutic compositions for treating allergic dermatitis that
CC comprise at least one ESP, as well as assay kits for testing if an
CC animal has, or is susceptible to, allergic dermatitis, and a method
CC of desensitising a host animal to allergic dermatitis using ESps.
CC The ESps can also be used for the production of antibodies useful
CC in diagnosis or in vaccines for passive immunisation against
CC allergic dermatitis.
XX
SQ Sequence 375 AA;
AAW82382 Length: 375 May 30, 2002 09:25 Type: P Check: 6654 ..
1 MEAILLTGIL ALTLSAECSE LKFVPMVKG PDHEACNYAG GPOLTTLOEK
51 DSVLTEDGKT EAYELGKLLD KYKKQQLKYD KWDATKTYWA VSTKAMRTKE
101 AALIVGAGLE NNPAAKAGNW TQOQLDSTHF DAMPGFSRFW NPQCCPAYFR
151 ALSLQNKIK KLEKYQTITI KEVTAKFPSI DGTKAQHIMI AYETFKRMKQ
201 QGRKEVEGIN TATMOKLKEF SSEFVLIALT STDOMRKLAG GLILKDLFND
251 IDELTKDHAQ PHAPGIRKN MNIFVVPQAI LAAQMAVEMP EGTKLRDQPI
301 TASNFYPDDQ SYVITIELYD KKNWNVQLQY KNNKNSGMLP IKVQGCNSPM

351 CPYDTLKSL NKYIIDARH KQACK

11AA_SEQUENCE 1.0

ID AAW82384 standard; Protein; 356 AA.

XX AAW82384;

DT 26-APR-1999 (first entry)

XX Flea saliva protein PfspN6-356.

XX Flea saliva protein; PfspN6-356; allergic dermatitis; allergen;

KM allergy; therapy; diagnosis; vaccine; ectoparasite.

XX Ctenocephalides sp.

PN W09845408-A2.

PD 15-OCT-1998.

PF 15-OCT-1997; 97WO-US18669.

PR 10-APR-1997; 97WO-US05959.

PA (HESK-) HESKA CORP.

PI Weber ER;

DR WPI; 1998-594480/50.

DR N-PSDB; AAV73434.

New isolated ectoparasite saliva genes - used to develop products for the diagnosis, prevention, treatment and determining susceptibility to allergic dermatitis

Claim 2; Page 137; 172pp; English.

This polypeptide comprises a new, claimed 40.5 kDa flea saliva protein, termed PfspN6-356, consisting of the mature polypeptide region of PfspN6-375 (see AAW82382). PfspN6-356 has been expressed in recombinant Escherichia coli BL-21 cells. The invention is directed to methods for isolating ectoparasite saliva proteins (ESPs). It provides ESps (claimed, see AAW82382-93), nucleic acid molecules encoding them, methods for their recombinant production, therapeutic compositions for treating allergic dermatitis that comprise at least one ESP, as well as assay kits for testing if an animal has, or is susceptible to, allergic dermatitis, and a method of desensitising a host animal to allergic dermatitis using ESps. The ESps can also be used for the production of antibodies useful in diagnosis or in vaccines for passive immunisation against allergic dermatitis.

SQ Sequence 356 AA;

AAW82384 Length: 356 May 30, 2002 09:25 Type: P Check: 2624 ..

1 ELKFEFVWVK GPDHEACNYA GGPQLTTLQE KDSVLTEDEK TEAYELGKLL

51 DKVYKKQLKV DKWDATKTYW AVSTKAMRTK EALIVGAGL ENNPAKAKGN

101 WTOOQLDSTH FDAMPGEFSRF WNPQCPAYF RALSLQNKI KLEKYQTT

151 IKEVTAKFPS IDGTAKQHIW IAYETFRMK QQGRKEVEGI NTATMOKLKE

201 FSSEFVLIAL TSTDQMRKLA GGLILKDLFN DIDELTKDHA OPHAPGIGKN

251 KMNIFVVPQA ILAAQMAVFM PEGTKLRDQP ITASNFYRDD QSYVILELYQ

301 DKNKWNVOLQ YKNNKSGWL PIKVQGCNSP MCPYDTLKSL NKYIIDARH

351 KQACK

11AA_SEQUENCE 1.0

ID AAW82385 standard; Protein; 355 AA.

XX AAW82385;

DT 26-APR-1999 (first entry)

XX Flea saliva protein PfspN6-357.

Flea saliva protein; PfspN6-357; allergic dermatitis; allergen; allergy; therapy; diagnosis; vaccine; ectoparasite.

XX Ctenocephalides sp.

PN W09845408-A2.

PD 15-OCT-1998.

PF 15-OCT-1997; 97WO-US18669.

PR 10-APR-1997; 97WO-US05959.

PA (HESK-) HESKA CORP.

PI Weber ER;

DR WPI; 1998-594480/50.

DR N-PSDB; AAV73436.

New isolated ectoparasite saliva genes - used to develop products for the diagnosis, prevention, treatment and determining susceptibility to allergic dermatitis

Example 10; Page 137; 172pp; English.

This polypeptide comprises the new, claimed 40.5 kDa flea saliva protein PfspN6-356 (see AAW82384) plus an additional N-terminal Met residue. It was expressed in Escherichia coli BL-21 cells transformed with a vector carrying a PCR amplification product (see AAV73436) termed PfspN6-1071. The invention is directed to products and methods for isolating ectoparasite saliva proteins (ESPs). It provides ESps (claimed, see AAW82382-93), nucleic acid molecules encoding them, methods for their recombinant production, therapeutic compositions for treating allergic dermatitis that comprise at least one ESP, as well as assay kits for testing if an animal has, or is susceptible to, allergic dermatitis, and a method of desensitising a host animal to allergic dermatitis using ESps. The ESps can also be used for the production of antibodies useful in diagnosis or in vaccines for passive immunisation against allergic dermatitis.

SQ Sequence 355 AA;

AAW82385 Length: 355 May 30, 2002 09:25 Type: P Check: 309 ..

1 MELKFEFVKG PDHEACNYAG GPQLTTLQEK DSVLTEDGKT EAYELGKLLD

51 KYKKQLKVD KWDATKTYWA VSTKAMRTKE AALIVGAGLE NNPAAKAGNW

101 TPOOQLDSTHF DAMPGFSRFW NPQCPAYFR ALSLQNKIK KLEKYQTTI

151 KEVTAKFPSI DGTAKQHIWI AYETFRMKQ QGRKEVEGIN TATMOKLKEE

201 SSEFVLIALT STDQMRKIAG GLILKDLFND IDELTAKDHAQ PHAPGIGKNK

251 MNIFVVPQAI ILAAQMAVFM PEGTKLRDQPI TASNFYRDDQ SYVILELYOD

301 KKNKWNVOLQ YKNNKSGWLP IKVQGCNSPM CPYDTLKSL NKYIIDARH

351 KQACK

11AA_SEQUENCE 1.0

ID AAW82322 standard; Peptide; 19 AA.

XX

```
AC AAW82322;
XX
DT 22-FEB-1999 (first entry)
XX
DE p53 homologue TIP 12/1 peptide.
XX
KW p53; mdm2; inhibitor; therapy; activator; treatment; cancer; medicament.
XX
OS Synthetic.
XX
PN WO9847919-A1.
XX
PD 29-OCT-1998.
XX
PF 20-APR-1998; 98WO-GB01140.
XX
PR 22-APR-1997; 97GB-0008089.
XX
PA (UYDU-) UNIV DUNDEE.
XX
PI Lane DP;
XX
DR WPI; 1998-609975/51.
XX
PT New substance with a mdm2 binding domain and coupling partner -
PT useful for stabilising in cells without an efficient mdm2-mediated
PT degradation pathway
XX
PS Disclosure; Fig 1; 52pp; English.
XX
CC This sequence is a peptide homologue of a region of p53 which binds
CC to mdm2. This peptide is used in the construction of a novel agent
CC capable of disrupting the binding of p53 and mdm2 or inhibiting the
CC production of mdm2 in a population of cells. This agent is also used in
CC the preparation of a therapeutic for activating p53, where the population
CC of cells do not overexpress mdm2. Inhibiting mdm2 production and/or
CC inhibiting the binding of p53 to p53 allows levels of p53 to increase by
CC reducing the clearance of p53 by mdm2, and can be used to activate p53
CC function. The agents for use in therapeutics for activating p53 can be
CC used for the treatment of cancer, viral conditions or other conditions
CC associated with non-functional p53.
XX
SQ Sequence 19 AA;

AAW82322 Length: 19 May 30, 2002 09:25 Type: P Check: 4400 ..
1 PPLSMPRFMD YWEGLENENG
11AA_SEQUENCE 1.0
ID AAW82320 standard; Peptide; 19 AA.
XX
AC AAW82320;
XX
DT 22-FEB-1999 (first entry)
XX
DE p53 homologue TIP 12/1 peptide.
XX
KW p53; mdm2; inhibitor; therapy; activator; treatment; cancer; medicament.
XX
OS Synthetic.
XX
PN WO9847525-A1.
XX
PD 29-OCT-1998.
XX
PF 20-APR-1998; 98WO-GB01144.
XX
PR 22-APR-1997; 97GB-0008092.
XX
PA (UYDU-) UNIV DUNDEE.
XX
PI Lane DP;
XX
```

```
DR WPI; 1998-609932/51.
XX
XX New agents which inhibit interaction of p53 and mdm2 - useful for
PT activating p53, e.g. for treating cancers, viral conditions or other
PT conditions associated with non functional p53 or mdm2
XX
PS Disclosure; Fig 1; 52pp; English.
XX
CC This sequence is a peptide homologue of a region of p53 which binds
CC to mdm2. This peptide is used in the construction of a novel agent
CC capable of disrupting the binding of p53 and mdm2 or inhibiting the
CC production of mdm2 in a population of cells. This agent is also used in
CC the preparation of a therapeutic for activating p53, where the population
CC of cells do not overexpress mdm2. Inhibiting mdm2 production and/or
CC inhibiting the binding of p53 to p53 allows levels of p53 to increase by
CC reducing the clearance of p53 by mdm2, and can be used to activate p53
CC function. The agents for use in therapeutics for activating p53 can be
CC used for the treatment of cancer, viral conditions or other conditions
CC associated with non-functional p53.
XX
SQ Sequence 19 AA;

AAW82320 Length: 19 May 30, 2002 09:25 Type: P Check: 4400 ..
1 PPLSMPRFMD YWEGLENENG
11AA_SEQUENCE 1.0
ID AAW58852 standard; Protein; 130 AA.
XX
AC AAW58852;
XX
DT 23-JUL-1998 (first entry)
XX
DE Human O289_1 secreted protein.
XX
KW Secreted protein; prevention; treatment; gene therapy.
XX
OS Homo sapiens.
XX
PN WO9801554-A2.
XX
PD 15-JAN-1998.
XX
PF 07-JUL-1997; 97WO-US11876.
XX
PR 09-JUL-1996; 96US-0677231.
XX
PA (GEMT ) GENETICS INST INC.
XX
PI Bowman M, Evans C, Jacobs K, Lavallic ER, McCoy JM;
PI Merberg D, Racie LA, Spaulding V, Treacy M;
XX
DR WPI; 1998-110230/10.
DR N-PSDB; AAV11438, AAV11439.
XX
PT Secreted proteins and polynucleotides encoding them - useful to
PT prevent, treat and ameliorate medical conditions
XX
PS Disclosure; Page 71; 93pp; English.
XX
CC This sequence represents a fragment of a novel secreted protein derived
CC from clone O289_1 which was isolated from a human dendritic cell cDNA
CC library. The protein can be used to prevent, treat or ameliorate a
CC medical condition, while the polynucleotides can be used for gene
CC therapy.
XX
SQ Sequence 130 AA;

AAW58852 Length: 130 May 30, 2002 09:25 Type: P Check: 5772 ..
1 NENPGVCHSS ATGDPHCROP EPTSPGPLSW LFSLPFPFR LAVTHRETSQ
51 LGVQDRDFRV FLSKEERDF YGSLPVGLLL AHRVDHTRKSC QFGFLSIPTP
```

101 SPRRVVTALI QALVISHLCH FSSFWIGLPA
11AA_SEQUENCE 1.0
ID AAW56025 standard; Protein; 316 AA.
XX
XX AAW56025;
XX
DT 28-JUL-1998 (first entry)
XX
DE Phosphate starvation-induced protein kinase psr6.
XX
KW Phosphate starvation-induced protein kinase; psr6; glucosidase;
KW phosphate transporter; phosphate deprivation; Arabidopsis thaliana;
KW Brassica nigra; photosynthetic organism; phosphorus metabolism; growth;
KW reproduction; metabolic content; flowering; drought; cold tolerance.
XX
OS Brassica nigra.
XX
PN WO9805760-A2.
XX
PD 12-FEB-1998.
XX
PF 30-JUL-1997; 97WO-CA00532.
XX
PR 31-JUL-1996; 96US-0688988.
PR 31-JUL-1996; 96CA-2182421.
XX
PA (TOOH) UNIV QUEENS KINGSTON.
XX
PI Lefebvre DD, Malboobi MA;
XX
DR WPI; 1998-159169/14.
DR N-PSDB; AAV26299.
XX
XX New isolated phosphate starvation-inducible genes - used to develop
PT plants or other photosynthetic organisms with altered phosphorus
PT metabolism, e.g. to alter growth or reproduction.
XX
XX Claim 62; Fig 21; 131pp; English.
XX
CC The present sequence represents a phosphate starvation-induced protein
CC kinase from a photosynthetic organism, in which transcription of the
CC DNA is induced by phosphate deficiency. The products and methods of
CC the present invention can be used to modify the phosphorus metabolism
CC in plants and other photosynthetic organisms, e.g. to alter growth,
CC reproduction, metabolic content, flowering, drought or cold tolerance
CC or nutritive value.
XX
SQ Sequence 316 AA;
AAW56025 Length: 316 May 30, 2002 09:25 Type: P Check: 9188 ..
1 IPGMTIMVMV LCSVASGLSF GDKPKSVMTT LCFERFWLGF GIGDYPPLSA
51 TIMSEYANKK TRGAFVSAVF AMQFGIMAG GIFAIISSA FEAKFPAPAY
101 AEDALASTVP QADFVWRIL MVGAIPAAMT YYSRSKMPET ARYTALVARD
151 AKQAASDMSR VLQVEIEAEQ EKVEIISNK SKAFSLFSKE FMKRHGLHL
201 GTTSTWFLD IAFYSQNLFQ KDIFSAIGWI PPAQTMNAIQ EVFKIARAQT
251 LIALGSTVPG YWFTVAFIDV IGRFAIQMG FEFMTVFEMFA LAIPYNHWT
301 TRNNLNCWKL KKRLND
11AA_SEQUENCE 1.0
ID AAW37220 standard; peptide; 12 AA.
XX
AC AAW37220;
XX
DT 20-JUL-1998 (first entry)

XX
DE MDM2 binding peptide unique phage insert sequence 1.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection; phage insert.
XX
OS Homo sapiens.
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Pickersley S;
XX
DR WPI; 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 8; Page 30; 45pp; English.
XX
CC This is a unique phage insert sequence of the MDM2 binding peptide
CC identified by phage display. The MDM2 binding peptides and their
CC derivatives are capable of binding to the human oncogenic protein MDM2.
CC These peptides can specifically inhibit or block the binding of MDM2 to
CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
CC peptides may be used to identify molecules that bind to MDM2 and to
CC identify and design inhibitors of MDM2/p53 binding. They may also be
CC used to purify binding partners especially MDM2, diagnose disease by
CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
CC for treatment or prevention of disease involving p53/MDM2 interactions,
CC especially tumours and viral infections. The peptides can be administered
CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
CC interaction, the peptides can activate p53 function and accumulation in
CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
CC a significantly greater blocking activity compared with wild-type p53.
XX
SQ Sequence 12 AA;
AAW37220 Length: 12 May 30, 2002 09:25 Type: P Check: 5978 ..
1 MPREMDYWEGLN
11AA_SEQUENCE 1.0
ID AAW37224 standard; peptide; 15 AA.
XX
AC AAW37224;
XX
DT 20-JUL-1998 (first entry)
XX
DE MDM2 binding peptide unique phage insert sequence 5.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection; phage insert.
XX
OS Homo sapiens.
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX


```
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 8; Page 30; 45pp; English.
XX
CC This is a unique phage insert sequence of the MDM2 binding peptide
CC identified by phage display. The MDM2 binding peptides and their
CC derivatives are capable of binding to the human oncogenic protein MDM2.
CC These peptides can specifically inhibit or block the binding of MDM2 to
CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
CC peptides may be used to identify molecules that bind to MDM2 and to
CC identify and design inhibitors of MDM2/p53 binding. They may also be
CC used to purify binding partners especially MDM2, diagnose disease by
CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
CC for treatment or prevention of disease involving p53/MDM2 interactions,
CC especially tumours and viral infections. The peptides can be administered
CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
CC interaction, the peptides can activate p53 function and accumulation in
CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
CC a significantly greater blocking activity compared with wild-type p53.
XX
SQ Sequence 15 AA;

AAW37224 Length: 15 May 30, 2002 09:25 Type: P Check: 9428 ..
1 PRPALVFADY WETLY

!1AA_SEQUENCE 1.0
ID AAW37221 standard; peptide: 12 AA.
XX
AC AAW37221;
XX
DT 20-JUL-1998 (first entry)
XX
DE MDM2 binding peptide unique phage insert sequence 2.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection; phage insert.
XX
OS Homo sapiens.
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI; 1998-100996/09.
```

```
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 8; Page 30; 45pp; English.
XX
CC This is a unique phage insert sequence of the MDM2 binding peptide
CC identified by phage display. The MDM2 binding peptides and their
CC derivatives are capable of binding to the human oncogenic protein MDM2.
CC These peptides can specifically inhibit or block the binding of MDM2 to
CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
CC peptides may be used to identify molecules that bind to MDM2 and to
CC identify and design inhibitors of MDM2/p53 binding. They may also be
CC used to purify binding partners especially MDM2, diagnose disease by
CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
CC for treatment or prevention of disease involving p53/MDM2 interactions,
CC especially tumours and viral infections. The peptides can be administered
CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
CC interaction, the peptides can activate p53 function and accumulation in
CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
CC a significantly greater blocking activity compared with wild-type p53.
XX
SQ Sequence 12 AA;
```

```
AAW37221 Length: 12 May 30, 2002 09:25 Type: P Check: 6151 ..
1 VQNFIDYWTQ QF

!1AA_SEQUENCE 1.0
ID AAW37222 standard; peptide: 12 AA.
XX
AC AAW37222;
XX
DT 20-JUL-1998 (first entry)
XX
DE MDM2 binding peptide unique phage insert sequence 3.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection; phage insert.
XX
OS Homo sapiens.
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 8; Page 30; 45pp; English.
XX
CC This is a unique phage insert sequence of the MDM2 binding peptide
CC identified by phage display. The MDM2 binding peptides and their
CC derivatives are capable of binding to the human oncogenic protein MDM2.
CC These peptides can specifically inhibit or block the binding of MDM2 to
CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
```


CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
CC peptides may be used to identify molecules that bind to MDM2 and to
CC identify and design inhibitors of MDM2/p53 binding. They may also be
CC used to purify binding partners especially MDM2, diagnose disease by
CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
CC for treatment or prevention of disease involving p53/MDM2 interactions,
CC especially tumours and viral infections. The peptides can be administered
CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
CC interaction, the peptides can activate p53 function and accumulation in
CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
CC a significantly greater blocking activity compared with wild-type p53.

XX Sequence 12 AA:

AAW37222 Length: 12 May 30, 2002 09:25 Type: P Check: 5993 ..

1 TGPFTHYWA TF

11AA_SEQUENCE 1.0

ID AAW37223 standard; peptide; 15 AA.

XX AAW37223;

DT 20-JUL-1998 (first entry)

DE MDM2 binding peptide unique phage insert sequence 4.

KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;

KW tumour; diagnosis; binding; viral infection; phage insert.

OS Homo sapiens.

PN WO9801467-A2.

PD 15-JAN-1998.

PF 04-JUL-1997; 97WO-EP03549.

PR 07-APR-1997; 97GB-0007041.

PR 05-JUL-1996; 96GB-0014197.

PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

PA (NOVS) NOVARTIS AG.

PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;

PI Hochkeppel H, Lane D, Picksley S;

DR WPI; 1998-100996/09.

PT Compounds binding to MDM2 protein and inhibit its interaction with

PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral

PT infections and identifying binding agents

XX Example 8; Page 30; 45pp; English.

XX This is a unique phage insert sequence of the MDM2 binding peptide
CC identified by phage display. The MDM2 binding peptides and their
CC derivatives are capable of binding to the human oncogenic protein MDM2.
CC These peptides can specifically inhibit or block the binding of MDM2 to
CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
CC peptides may be used to identify molecules that bind to MDM2 and to
CC identify and design inhibitors of MDM2/p53 binding. They may also be
CC used to purify binding partners especially MDM2, diagnose disease by
CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
CC for treatment or prevention of disease involving p53/MDM2 interactions,
CC especially tumours and viral infections. The peptides can be administered
CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
CC interaction, the peptides can activate p53 function and accumulation in
CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
CC a significantly greater blocking activity compared with wild-type p53.

XX Sequence 15 AA:

AAW37223 Length: 15 May 30, 2002 09:25 Type: P Check: 9093 ..

1 IDRPTFRDH WFALV

11AA_SEQUENCE 1.0

ID AAW37225 standard; peptide; 15 AA.

XX AAW37225;

DT 20-JUL-1998 (first entry)

DE MDM2 binding peptide unique phage insert sequence 6.

KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;

KW tumour; diagnosis; binding; viral infection; phage insert.

OS Homo sapiens.

PN WO9801467-A2.

PD 15-JAN-1998.

PF 04-JUL-1997; 97WO-EP03549.

PR 07-APR-1997; 97GB-0007041.

PR 05-JUL-1996; 96GB-0014197.

PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

PA (NOVS) NOVARTIS AG.

PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;

PI Hochkeppel H, Lane D, Picksley S;

DR WPI; 1998-100996/09.

PT Compounds binding to MDM2 protein and inhibit its interaction with

PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral

PT infections and identifying binding agents

XX Example 8; Page 30; 45pp; English.

XX This is a unique phage insert sequence of the MDM2 binding peptide
CC identified by phage display. The MDM2 binding peptides and their
CC derivatives are capable of binding to the human oncogenic protein MDM2.
CC These peptides can specifically inhibit or block the binding of MDM2 to
CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
CC peptides may be used to identify molecules that bind to MDM2 and to
CC identify and design inhibitors of MDM2/p53 binding. They may also be
CC used to purify binding partners especially MDM2, diagnose disease by
CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
CC for treatment or prevention of disease involving p53/MDM2 interactions,
CC especially tumours and viral infections. The peptides can be administered
CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
CC interaction, the peptides can activate p53 function and accumulation in
CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
CC a significantly greater blocking activity compared with wild-type p53.

XX Sequence 15 AA:

AAW37225 Length: 15 May 30, 2002 09:25 Type: P Check: 8833 ..

1 PAFSRFWSDL SAGAH

11AA_SEQUENCE 1.0

ID AAW37226 standard; peptide; 10 AA.

XX AAW37226;

XX

```
DT 20-JUL-1998 (first entry)
XX
XX MDM2 binding peptide phage consensus sequence.
DE
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KM tumour; diagnosis; binding; viral infection; phage insert.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 2
FT Misc-difference 2 /note= "any natural amino acid"
FT Misc-difference 4
FT Misc-difference 4 /note= "any natural amino acid"
FT Misc-difference 8
FT Misc-difference 8 /note= "any natural amino acid"
FT Misc-difference 9
FT Misc-difference 9 /note= "any natural amino acid"
XX
XX WO9801467-A2.
PN
XX 15-JAN-1998.
PD
XX
XX 04-JUL-1997; 97WO-EP03549.
PF
XX
XX 07-APR-1997; 97GB-0007041.
PR
XX 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
XX WPI; 1998-100996/09.
DR
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Claim 5; Page 41; 45pp; English.
XX
XX This is a phage consensus sequence of the MDM2 binding peptide identified
CC by phage display. This shows similarity to the known MDM2 binding motif
CC on p53 (TFSDLM). The MDM2 binding peptides and their derivatives are
CC capable of binding to the human oncogenic protein MDM2. These peptides
CC can specifically inhibit or block the binding of MDM2 to the human p53
CC protein, in vitro or in vivo. Inhibiting the interaction between the p53
CC and MDM2 can induce growth arrest or apoptosis in tumour cells comprising
CC a wild-type p53 and non-elevated levels of MDM2. The peptides may be used
CC to identify molecules that bind to MDM2 and to identify and design
CC inhibitors of MDM2/p53 binding. They may also be used to purify binding
CC partners especially MDM2, diagnose disease by measuring levels of MDM2 in
CC blood of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally or
CC by injection. By interfering with MDM2/p53 interaction, the peptides can
CC activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
XX Sequence 10 AA;
SQ
AAW37226 Length: 10 May 30, 2002 09:25 Type: P Check: 4557 ..
1 PFXDYMXXL
1!AA_SEQUENCE 1.0
ID AAW37195 standard; peptide; 12 AA.
XX
XX AAW37195;
AC
XX 20-JUL-1998 (first entry)
DT
```

```
XX
XX Human oncogenic protein MDM2 binding C-amidated peptide derivative 1.
DE
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KM tumour; diagnosis; binding; viral infection.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 12
FT Modified-site 12 /note= "C-terminal amide"
FT
XX
XX WO9801467-A2.
PN
XX 15-JAN-1998.
PD
XX
XX 04-JUL-1997; 97WO-EP03549.
PF
XX
XX 07-APR-1997; 97GB-0007041.
PR
XX 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
XX WPI; 1998-100996/09.
DR
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Example 1; Page 20; 45pp; English.
XX
XX This is a C-amidated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
XX Sequence 12 AA;
SQ
AAW37195 Length: 12 May 30, 2002 09:25 Type: P Check: 5993 ..
1 TGPATHYWA TF
1!AA_SEQUENCE 1.0
ID AAW37196 standard; peptide; 12 AA.
XX
XX AAW37196;
AC
XX 20-JUL-1998 (first entry)
DT
XX
XX Human oncogenic protein MDM2 binding C-amidated peptide derivative 2.
DE
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KM tumour; diagnosis; binding; viral infection.
XX
XX Synthetic.
OS Homo sapiens.
OS
```

```
XX Key Location/Qualifiers
FH Modified-site 12
FT /note= "C-terminal amide"
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
XX WPI; 1998-100996/09.
DR
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Example 1; Page 20; 45pp; English.
PS
XX This is a C-amidated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
XX Sequence 12 AA;
SQ
AAW37196 Length: 12 May 30, 2002 09:25 Type: P Check: 5978 ..
1 MPRFMDYMEG LN
11AA_SEQUENCE 1.0
ID AAW37197 standard; peptide; 14 AA.
XX
XX AAW37197;
AC
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding Cys (Acrid) peptide derivative 1.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "N-terminal acetyl; Sulphydryl side-chain
FT linked to 6-acryloyl-2-(dimethylamino)
FT naphthalene"
FT Modified-site 14
FT /note= "C-terminal amide"
```

```
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
XX WPI; 1998-100996/09.
DR
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Example 2; Page 20; 45pp; English.
PS
XX This is a Cys (Acrid) peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
XX Sequence 14 AA;
SQ
AAW37197 Length: 14 May 30, 2002 09:25 Type: P Check: 8253 ..
1 CGQPTFSDYW KLLP
11AA_SEQUENCE 1.0
ID AAW37203 standard; peptide; 16 AA.
XX
XX AAW37203;
AC
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding biotinylated peptide derivative 3.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection; biotinylation.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "biotinylated"
FT Modified-site 16
FT /note= "C-terminal amide"
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX
```

```
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 3; Page 21; 45pp; English.
XX
XX This is a biotinylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 16 AA;
AAW37203 Length: 16 May 30, 2002 09:25 Type: P Check: 634 ..
1 SSGQETFFSD YMKLLP
!!AA_SEQUENCE 1.0
ID AAW37204 standard; peptide; 28 AA.
XX
AC AAW37204;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding biotinylated peptide derivative 4.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection; biotinylation.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "biotinylated"
FT Modified-site 28 /note= "C-terminal amide"
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
```

```
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 3; Page 21; 45pp; English.
XX
XX This is a biotinylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 28 AA;
AAW37204 Length: 28 May 30, 2002 09:25 Type: P Check: 1571 ..
1 SMPRFMDYWE GLNRQIKIWF QNRMKWK
!!AA_SEQUENCE 1.0
ID AAW37205 standard; peptide; 31 AA.
XX
AC AAW37205;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding biotinylated peptide derivative 5.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection; biotinylation.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 17 /label= bAla
FT Modified-site 30 /note= "beta-Alanine"
FT Modified-site 31 /label= bAla
FT Modified-site 31 /note= "beta-Alanine"
FT Modified-site 31 /note= "biotinylated"
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
```

XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI: 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 3; Page 21; 45pp; English.
XX
CC This is a biotinylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 31 AA;
AAW37205 Length: 31 May 30, 2002 09:25 Type: P Check: 8233 ..
1 AAVALLPAVL LALLAPXMPR FMDYWGGLNX K
11AA_SEQUENCE 1.0
ID AAW37182 standard; peptide; 12 AA.
XX
AC AAW37182;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 1.
XX
DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 1.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 12 /note= "C-terminal amide"
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX
PE 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI: 1998-100996/09.
XX

PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 1; Page 18; 45pp; English.
XX
CC This is a N-acylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 12 AA;
AAW37182 Length: 12 May 30, 2002 09:25 Type: P Check: 5993 ..
1 TGPAFTHYWA TF
11AA_SEQUENCE 1.0
ID AAW37183 standard; peptide; 12 AA.
XX
AC AAW37183;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 2.
XX
DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 2.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 12 /note= "C-terminal amide"
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX
PE 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI: 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 1; Page 19; 45pp; English.
XX

CC This is a N-acylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.

XX Sequence 12 AA;

AAW37183 Length: 12 May 30, 2002 09:25 Type: P Check: 5978 ..

1 MPRFMDYWEGLN

1!AA_SEQUENCE 1.0

ID AAW37184 standard; peptide; 12 AA.

XX AAW37184;

DT 20-JUL-1998 (first entry)

DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 3.

KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;

KM tumour; diagnosis; binding; viral infection.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 12 /note= "C-terminal amide"

PN WO9801467-A2.

XX 15-JAN-1998.

PF 04-JUL-1997; 97WO-EP03549.

PR 07-APR-1997; 97GB-0007041.

PR 05-JUL-1996; 96GB-0014197.

PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

PA (NOVS) NOVARTIS AG.

PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;

PI Hochkeppel H, Lane D, Picksley S;

DR WPI; 1998-100996/09.

PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents

PS Example 1; Page 19; 45pp; English.

XX This is a N-acylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify

CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.

XX Sequence 12 AA;

AAW37184 Length: 12 May 30, 2002 09:25 Type: P Check: 6146 ..

1 QPTFSDYWKLLP

1!AA_SEQUENCE 1.0

ID AAW37190 standard; peptide; 12 AA.

XX AAW37190;

DT 20-JUL-1998 (first entry)

DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 9.

KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;

KM tumour; diagnosis; binding; viral infection.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 12 /note= "C-terminal amide"

PN WO9801467-A2.

XX 15-JAN-1998.

PF 04-JUL-1997; 97WO-EP03549.

PR 07-APR-1997; 97GB-0007041.

PR 05-JUL-1996; 96GB-0014197.

PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

PA (NOVS) NOVARTIS AG.

PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;

PI Hochkeppel H, Lane D, Picksley S;

DR WPI; 1998-100996/09.

PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents

PS Example 1; Page 19; 45pp; English.

XX This is a N-acylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally

CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 12 AA;

AAW37190 Length: 12 May 30, 2002 09:25 Type: P Check: 6124 ..

1 QETFSDYWKL LP

!!AA_SEQUENCE 1.0

ID AAW37191 standard; peptide; 12 AA.

XX AAW37191;

DT 20-JUL-1998 (first entry)

DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 10.

XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;

KW tumour; diagnosis; binding; viral infection.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 1 /note- "N-terminal acetyl"

FT Modified-site 12 /note- "C-terminal amide"

FT

XX WO9801467-A2.

PD 15-JAN-1998.

XX 04-JUL-1997; 97WO-EP03549.

XX 07-APR-1997; 97GB-0007041.

PR 05-JUL-1996; 96GB-0014197.

XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

PA (NOVS) NOVARTIS AG.

XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;

PI Hochkeppel H, Lane D, Picksley S;

XX WPI; 1998-100996/09.

XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents

XX Example 1; Page 19; 45pp; English.

XX This is a N-acylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.

SQ Sequence 12 AA;

AAW37191 Length: 12 May 30, 2002 09:25 Type: P Check: 6151 ..
1 VQNFIDYWTQ QF

!!AA_SEQUENCE 1.0

ID AAW37192 standard; peptide; 15 AA.

XX AAW37192;

DT 20-JUL-1998 (first entry)

DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 11.

XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;

KW tumour; diagnosis; binding; viral infection.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 1 /note- "N-terminal acetyl"

FT Modified-site 15 /note- "C-terminal amide"

FT

XX WO9801467-A2.

PD 15-JAN-1998.

XX 04-JUL-1997; 97WO-EP03549.

XX 07-APR-1997; 97GB-0007041.

PR 05-JUL-1996; 96GB-0014197.

XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

PA (NOVS) NOVARTIS AG.

XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;

PI Hochkeppel H, Lane D, Picksley S;

XX WPI; 1998-100996/09.

XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents

XX Example 1; Page 19; 45pp; English.

XX This is a N-acylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.

XX Sequence 15 AA;

AAW37192 Length: 15 May 30, 2002 09:25 Type: P Check: 9093 ..

1 IDRAPTFRDH WFALV

!!AA_SEQUENCE 1.0

ID AAW37193 standard; peptide; 15 AA.
XX
AC AAW37193;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 12.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX
KM tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 15
FT /note= "C-terminal amide"
XX
PN W09801467-A2.
XX
PD 15-JAN-1998.
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 1; Page 19; 45pp; English.
XX
CC This is a N-acylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 15 AA;
AAW37193 Length: 15 May 30, 2002 09:25 Type: P Check: 9428 ..
1 PRPALVFADY WETLY
!!AA_SEQUENCE 1.0
ID AAW37194 standard; peptide; 15 AA.
XX
AC AAW37194;
XX
DT 20-JUL-1998 (first entry)
XX

DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 13.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX
KM tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 15
FT /note= "C-terminal amide"
XX
PN W09801467-A2.
XX
PD 15-JAN-1998.
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 1; Page 19; 45pp; English.
XX
CC This is a N-acylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 15 AA;
AAW37194 Length: 15 May 30, 2002 09:25 Type: P Check: 8833 ..
1 PAFSRFWSDL SAGAH
!!AA_SEQUENCE 1.0
ID AAW37170 standard; peptide; 12 AA.
XX
AC AAW37170;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding peptide 1.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX
KM tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.

OS Homo sapiens.
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Claim 5; Page 41; 45pp; English.
XX
CC This peptide is capable of binding to an oncogenic protein MDM2
CC (especially human DM2). The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 12 AA;
AAW37170 Length: 12 May 30, 2002 09:25 Type: P Check: 5978 ..
1 MPRFMDYWEGLN
11AA_SEQUENCE 1.0
ID AAW37171 standard; peptide; 12 AA.
XX
AC AAW37171;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding peptide 2.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Claim 5; Page 41; 45pp; English.
XX
CC This peptide is capable of binding to an oncogenic protein MDM2
CC (especially human DM2). The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 12 AA;
AAW37171 Length: 12 May 30, 2002 09:25 Type: P Check: 6146 ..
1 QPTFSDYWKLLP
11AA_SEQUENCE 1.0
ID AAW37172 standard; peptide; 15 AA.
XX
AC AAW37172;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding peptide 3.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
DR WPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Disclosure; Page 4; 45pp; English.

XX This peptide is capable of binding to an oncogenic protein MDM2
CC (especially human DM2). The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.

XX Sequence 15 AA;

AAW37172 Length: 15 May 30, 2002 09:25 Type: P Check: 9428 ..

1 PRPALVFADY WETLY

11AA_SEQUENCE 1.0

ID AAW46748 standard; Protein; 534 AA.

AC AAW46748;

DT 06-JUL-1998 (first entry)

DE Arabidopsis thaliana phosphate transporter 2 protein.

KW Phosphate transporter protein; AtPT1; AtPT2; LePT1; LePT2;
KW phosphate-deficiency.

OS Arabidopsis thaliana.

OS Key Location/Qualifiers

FT Misc-difference 290

FT /note- "encoded by TTC"

FT Misc-difference 437

FT /note- "encoded by TTC"

PN WO9804701-A1.

PD 05-FEB-1998.

PF 29-JUL-1997; 97WO-US13458.

PR 29-JUL-1996; 96US-0022391.

PA (PURD) PURDUE RES FOUND.

PI Muchhal US, Raghothama KG;

DR WPI; 1998-130690/12.

DR N-PSDB; AAV16291.

PT Plant phosphate transporter proteins - used to improve plants

PT ability to grow in phosphate-deficient soil

PS Claim 2; Page 38-40; 66pp; English.

XX This sequence is the Arabidopsis thaliana phosphate transporter
CC protein 2 (AtPT2). AtPT1 (see AAW46747) AtPT2 (see AAW46748), LePT1 (see
CC AAW46749), and LePT2 (see AAW46750) are the proteins of the invention.
CC AtPT1 and AtPT2 are from Arabidopsis thaliana, while LePT1 and LePT2 are
CC the tomato equivalents of the Arabidopsis proteins. The DNA sequences,
CC and constructs and vectors containing them can be used for improving a
CC plant's ability to grow in phosphate-deficient soil.

XX Sequence 534 AA;

AAW46748 Length: 534 May 30, 2002 09:25 Type: P Check: 562 ..
1 MAREQLQVLN ALDVAKTQWY HFTAILIAGM GFFTDAYDLF CISLVTKLLG
51 RIYHVEGAQ KPGTLPNVA AAVNGVAFCG TLAGQLFFGW LGDKLGRKKV
101 YGNTLMVMVL CSTASGLSFG HEPRKVMATL CFFRFWLFG IGGDYPLSAT
151 IMSEYANKKT RCAFVSAVFA MCGFGIMAGC IFAILISSAF EAKFPPPAYA
201 DDLGSTITPQ ADLWRIITM AGAIPAAWY YSRSKMPETA RTALVAKDA
251 KQASDMSKV LOVEIEPEQO KLEISKESK KAFGLFSKEE MSRHGLHLTG
301 TTSTWFLLDI AFYSQNLFOK DIFSAIGWIP PAQSMNAIQE VFKIARAQTL
351 TALCSTVPGY WFTVAFIDVI GRFAIQMGF FFWTFMFAL AIPYNHWTIK
401 ENRIGFVIMY SLTEFFANFG PNATTFVPA EIPPARERST CHGISAASGK
451 LGAMVGAFGF LYLAQNPKDK KTDAGYPPGI GVRNSLIYLG VVNFGLILFT
501 FLVPESSGKS LEEMSGENED NENSNDSTRT VPIV

11AA_SEQUENCE 1.0

ID AAW46749 standard; Protein; 538 AA.

AC AAW46749;

DT 06-JUL-1998 (first entry)

DE Tomato phosphate transporter 1 protein.

KW Phosphate transporter protein; AtPT1; AtPT2; LePT1; LePT2; tomato;
KW phosphate-deficiency.

OS Lycopersicon esculentum.

PN WO9804701-A1.

PD 05-FEB-1998.

PF 29-JUL-1997; 97WO-US13458.

PR 29-JUL-1996; 96US-0022391.

PA (PURD) PURDUE RES FOUND.

PI Muchhal US, Raghothama KG;

DR WPI; 1998-130690/12.

DR N-PSDB; AAV16292.

PT Plant phosphate transporter proteins - used to improve plants

PT ability to grow in phosphate-deficient soil

PS Claim 2; Page 41-44; 66pp; English.

XX This sequence is the Lycopersicon esculentum (tomato) phosphate
CC transporter protein 1 (LePT1). AtPT1 (see AAW46747) AtPT2 (see
CC AAW46748), LePT1 (see AAW46749), and LePT2 (see AAW46750) are the
CC proteins of the invention. AtPT1 and AtPT2 are from Arabidopsis thaliana,
CC while LePT1 and LePT2 are the tomato equivalents of the Arabidopsis
CC proteins. The DNA sequences, and constructs and vectors containing them
CC can be used for improving a plant's ability to grow in
CC phosphate-deficient soil.

XX Sequence 538 AA;

AAW46749 Length: 538 May 30, 2002 09:25 Type: P Check: 4793 ..

1 MANDLOVLNA LDVAKTQLYH FTAIVIAGMG FFTDAYDLFC ISMVKLLIGR

51 LYHHHDGALK PGSLPPNVSA AVNGVAFCGT LAGQLEFGWL GDKMGRKKVY
 101 GMTLMIMVIC SIASGLSEGH TPKGVMITLC FFRFWLGFGI GGDYPLSATI
 151 MSEYANKKTR GAFIAAVFAM OGEGILAGCM VALIISAFAK GAFPAPAYEV
 201 DAIGSTVPPQ DEYVRIITLMF GAIPAGLTYT WRMKMPETAR YTALVAKNLK
 251 QAANDMSKVL QVEIEAPEK VTAISEAKGA NDFGLFTKEF LRRHGLHLIG
 301 TASTWFLLDI AEXSONLFQK DIFSAIGWIP PAQTMALAE VYKIAQAQTL
 351 IALCSTVPGY WETVAFIDKI GRPAIQLMGF FEMTVEMFAL ALPYHHWTLK
 401 DHRIGFVVMY SEFTFFANFG PNATTFVPA EIFPARLRST CHGISAAAGK
 451 AGAMVGAFFG LYAAQPTDPT KTDAGYPPGI GVRNSLIVLG CVNELGMLFT
 501 FLVPESSNGKS LEDLSRENEG EETVAEIRA TSGRTVPV

11AA_SEQUENCE 1.0
 ID AAW46750 standard; Protein; 528 AA.

AC AAW46750;
 DT 06-JUL-1998 (first entry)
 DE Tomato phosphate transporter 1 protein.

KW Phosphate transporter protein; AtPT1; AtPT2; LePT1; LePT2; tomato;
 KW phosphate-deficiency.

OS Lycopersicon esculentum.

Key Location/Qualifiers
 FT Misc-difference 364 /note="encoded by GCG"

PN WO9804701-A1.

PD 05-FEB-1998.

PF 29-JUL-1997; 97WO-US13458.

PR 29-JUL-1996; 96US-0022391.

PA (PURD) PURDUE RES FOUND.

PI Muchhal US, Raghochama KG;

DR WPI; 1998-130690/12.

DR N-PSDB; AAV16293.

PT Plant phosphate transporter proteins - used to improve plants
 PT ability to grow in phosphate-deficient soil

PS Claim 2; Page 41-44; 66pp; English.

CC This sequence is the Lycopersicon esculentum (tomato) phosphate
 CC transporter protein 1 (LePT1). AtPT1 (see AAW46747) AtPT2 (see
 CC AAW46748), LePT1 (see AAW46749), and LePT2 (see AAW46750) are the
 CC proteins of the invention. AtPT1 and AtPT2 are from Arabidopsis thaliana,
 CC while LePT1 and LePT2 are the tomato equivalents of the Arabidopsis
 CC proteins. The DNA sequences, and constructs and vectors containing them
 CC can be used for improving a plant's ability to grow in
 CC phosphate-deficient soil.

XX Sequence 528 AA;

AAW46750 Length: 528 May 30, 2002 09:25 Type: P Check: 3154 ..

1 MAVGDNDNNN IQVINALDLA KTQLYHFTAI VIAGMGFTTD AYDLFSISLV

51 TKLLGRLYYT KPDLKPGTL PPAVSASVTG VALVGTLAGQ LFGWLGDKM
 101 GRKKYGMTL VLMVCSVAS GLSEGSTPKG VMITLCEFRF WLGEGIGDGY
 151 PLSATIMSEY ANKRTGAFI AAVFAMOGFG ILFSGIVALI TAAGFDHAYR
 201 SPPEENNAL STVPQSDIYW RIIMFGSLP AALTYWRMK MBETARYTAL
 251 VAKDAKRAAQ DMGVLQVEI ESEBAKIEQI SRNETNQFGL FSWEFVRRHG
 301 LHLFGTCSTW FLIDIAFYSQ NLFQKDVESA VGMIPKAPTM NAVQEVYKIA
 351 RAQTLIALCS TVPCWFTVA FIDIIGRFPI QMGFEFMTV FMPAIAIPYH
 401 HWTLEANRIG FIVMYSLTFF FANEGPNATF FVPAEIPFA RLSTCHGIS
 451 AAAGKAGAIY GAYGFLYAAQ SKDPNKTDAG YPAGIGIKNS LIVGCINAL
 501 GMLCTFCVPE PKGKSLEAS QETITGEA

11AA_SEQUENCE 1.0
 ID AAW37145 standard; Protein; 130 AA.

AC AAW37145;
 DT 18-JUN-1998 (first entry)
 DE Human O289_1 protein.

DE Human O289_1 protein.

KW Human; O289_1; secreted protein; immunomodulator; antitumour;
 KW tissue growth; haemostatic; thrombolytic; cancer.

OS Homo sapiens.

PN WO9801552-A2.

PD 15-JAN-1998.

PF 07-JUL-1997; 97WO-US11842.

PR 27-SEP-1996; 96US-0721488.

PR 09-JUL-1996; 96US-0677231.

PA (GEM) GENETICS INST INC.

PI Bowman M, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racie LA, Spaulding, Treacy M;

DR WPI; 1998-101050/09.

DR N-PSDB; AAV02990.

PT New nucleic acid encoding secreted human proteins, and related
 PT vectors, transformed cells and proteins - are potentially useful
 PT e.g. as immuno-modulators, antitumour agents, promoters of tissue
 PT growth, haemostatic and thrombolytic agents etc.

PS Claim 17; Page 45; 55pp; English.

CC The present sequence represents the human O289_1 protein, which
 CC is a secreted protein. The DNA encoding the secreted protein may be used
 CC e.g. to express the recombinant protein; as tissue/molecular weight
 CC marker; for chromosome identification; to identify possible genetic
 CC disorders; to isolate new related DNA; as sources of primers for
 CC polymerase chain reaction; to generate anti-protein or anti-DNA
 CC antibodies or in interaction trap assays to identify sequences that
 CC encode interacting proteins. The protein can be used to screen compounds
 CC for biological activity; to raise antibodies; as tissue markers; for
 CC isolation of related receptors and ligands and as nutritional sources.
 CC It may also have many biological activities, e.g. cytokine and cell
 CC proliferation/differentiation activity; immunosuppressant/immunostimulant
 CC activity; regulation of haematopoiesis; inhibition or activation of
 CC follicle-stimulating hormone; chemotactic/chemokinetic activity;

CC haemostatic and thrombolytic activity; as receptors or ligands;
CC antiinflammatory activity; as antimicrobials; modulators of metabolism or
CC behaviour; as analgesics; enzymes for treating specific deficiency
CC disorders, and in treatment of psoriasis. Neutralising antibodies
CC against the protein can be used therapeutically, e.g. to detect or
CC prevent metastasis of the protein-expressing cancers.

XX Sequence 130 AA;

AAW37145 Length: 130 May 30, 2002 09:25 Type: P Check: 5772 ..

1 NENPGVCHSS ATGDPHCRQP EPTSPGLSM LFSLLPFPPR LAVTHRETQ

51 LGVQDRDFRV FLSKERDDE YGSLPVGILL AHRVDHTKSC QFGFLSIFTP

101 SPRVRVTALI QALVISHLCH FSSFWIGLPA

11AA_SEQUENCE 1.0

ID AAY30723 standard; Protein; 182 AA.

XX AC AAY30723;

XX DT 18-NOV-1999 (first entry)

XX DE Amino acid sequence of a human secreted protein.

XX Secreted protein; gene therapy; cancer; tumor; fetal deficiency;
KW neurodegenerative disorder; developmental abnormality; blood disorder;
KW immune system disease; autoimmune disease; leukemia; inflammation;
KW allergy; Alzheimer's disease; cognitive disorder; schizophrenia;
KW obesity; osteoporosis; arthritis; infection; AIDS; diabetes; asthma;
KW connective tissue disorder; transplant rejection; sepsis; acne;
KW psoriasis; cardiovascular disorder; reproductive disorder;
KW food additive; food preservative; storage capability.

XX OS Homo sapiens.

XX PN WO9943693-A1.

XX PD 02-SEP-1999.

XX PF 24-FEB-1999; 99WO-US03939.

XX PR 26-FEB-1998; 98US-0076051.

XX PR 26-FEB-1998; 98US-0076052.

XX PR 26-FEB-1998; 98US-0076053.

XX PR 26-FEB-1998; 98US-0076054.

XX PR 26-FEB-1998; 98US-0076057.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Olsen HS, Florence K, Brewer LA, Ebner R, Ruben SM, Rosen CA;

XX PI Duan RD;

XX DR WPI; 1999-550857/46.

XX DR N-PSDB; AAZ10662.

XX PT New human genes and the secreted polypeptides they encode, useful for

XX PT diagnosis and treatment of e.g. cancers, neurological disorders, immune

XX PT diseases, inflammation or blood disorders

XX PS Claim 11; Page 201; 246pp; English.

XX AAY30701-37 represent human secreted proteins. The polynucleotides and
CC their corresponding secreted polypeptides are useful for preventing,
CC treating or ameliorating medical conditions, e.g. by protein or gene
CC therapy. Pathological conditions can also be diagnosed by determining
CC the amount of the new polypeptides in a sample or by determining the
CC presence of mutations in the polynucleotide. Specific uses include
CC developing products for the diagnosis or treatment of cancer, tumors,
CC neurodegenerative disorders, developmental abnormalities and fetal
CC deficiencies, blood disorders, sepsis, diseases of the immune
CC system, autoimmune diseases, inflammation, allergies, Alzheimer's and

CC cognitive disorders, schizophrenia, obesity, osteoporosis, arthritis,
CC infections, AIDS, connective tissue disorders, transplant rejection,
CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
CC and reproductive disorders. The polypeptides or polynucleotides can
CC also be used as food additives or preservatives, such as to increase
CC or decrease storage capabilities, fat content, lipid, protein,
CC carbohydrate, vitamins, minerals, cofactors or other nutritional
CC components.

XX Sequence 182 AA;

AAY30723 Length: 182 May 30, 2002 09:25 Type: P Check: 7459 ..

1 MTSRRSSTLS MTSSLSLGLC ALTSAPAST MSWPLQLML DQSPRRVARK

51 SVSQLCPLLR PHPLSSKHP LVLPLQLPPT FLHLLPGPGC PGQTVAYWLL

101 EFLSRATLKL YPGDRPLWLQ PTRLNFKDHW TIFSVASAL FCVHRMATDR

151 HASFPTHWKA HRQGERGHRR CQHCYSKDL KX

11AA_SEQUENCE 1.0

ID AAY31622 standard; Protein; 284 AA.

XX AC AAY31622;

XX DT 02-NOV-1999 (first entry)

XX DE Human oxidised LDL receptor (HOLR) cDNA.

XX Human oxidised LDL receptor; HOLR; low density lipoprotein; LDL;
KW cardiovascular disorder; atherosclerosis; hypertension; stroke;
KW cancer; autoimmune disorder; diagnosis; treatment.

XX OS Homo sapiens.

XX FH Key

XX FT Modified-site

XX FT Modified-site

XX FT Region

XX FT Domain

XX FT Misc-difference

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX PN US5945308-A.

```

XX 31-AUG-1999.
PD
XX
XX 03-APR-1998; 98US-0055095.
PF
XX
XX 03-APR-1998; 98US-0055095.
PR
XX
XX (INCY-) INCYTE PHARM INC.
PA
XX
XX Corley MC, Patterson C, Sather S, Tang YT;
PI
XX WPI; 1999-508172/42.
DR
XX N-PSDB; AA211726.
DR
XX
XX Isolated and purified polynucleotide encoding an LDL receptor
PT
XX
XX Claim 1; Fig 1; 31pp; English.
PS
XX
XX This sequence represents a human oxidised LDL receptor (HOLR).
CC
XX cDNA sequences encoding HOLR were first identified in a clone from a
CC foetal lung cDNA library, and a consensus sequence produced from several
CC overlapping and/or extended nucleic acid sequences. The cDNA, and hence
CC this sequence are therefore consensus. High plasma concentrations of
CC LDL and low plasma concentrations of HDL (high density lipoprotein) are
CC associated with atherosclerosis and hypertension. Some of the
CC deleterious effects of LDL may be associated with an oxidised derivative
CC of LDL, and its uptake and subsequent degradation in vascular
CC endothelial cells as been implicated in the genesis of atherosclerotic
CC plaques. The HOLR cDNA and sequences complementary to it may be used
CC for the diagnosis and/or treatment (antisense therapy) of disorders
CC associated with the expression of HOLR, including cardiovascular
CC disorders (e.g., atherosclerosis, hypertension, stroke), various cancers
CC and autoimmune disorders (e.g., AIDS, diabetes mellitus, multiple
CC sclerosis, rheumatoid arthritis). HOLR protein antagonists may also be
CC used to treat such disorders.
XX
XX Sequence 284 AA;
SQ
AAY31622 Length: 284 May 30, 2002 09:25 Type: P Check: 7787
..
1 MLDDDGDTTM SLHSQASATA RHEPRRTEH RAPSSTWRPV ALTLLTCLV
51 LLIGLALGL LCKSALXPGG GSWFQVFOY QLSNTGQDTI SQMEERLGN
101 SOELSLQVQ NIKLAGSLQH VAEKLCRELY NKAGAHRCSP CTEQWKWHD
151 NCYQFYKDSK SWEDCKYFCL SENSTMLKIN KQEDLEFAAS QSYSEFFYSY
201 WTGLLRPDSG KAWLWMDGTP FTSELFHIII DVTSPRRDC VAILNGMIFS
251 KDCCKELKRCV CERRAGMYKP ESLHVPETL GEGD
!!AA_SEQUENCE 1.0
ID AAY38483 standard; Protein; 86 AA.
XX
XX AAY38483;
AC
XX 18-OCT-1999 (first entry)
DT
XX
XX Human secreted protein encoded by gene No. 24.
DE
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
OS
XX
XX WO9935158-A1.
PN

```

```

XX 15-JUL-1999.
PD
XX
XX 06-JAN-1999; 99WO-US00108.
PF
XX
XX 07-JAN-1998; 98US-0070704.
PR
XX 07-JAN-1998; 98US-0070657.
PR
XX 07-JAN-1998; 98US-0070658.
PR
XX 07-JAN-1998; 98US-0070692.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;
PI
XX Olsen HS, Rosen CA, Ruben SM, Soppet DR;
PI
XX
XX WPI; 1999-444190/37.
DR
XX N-PSDB; AA206242.
DR
XX
XX New isolated human genes and the secreted polypeptides they encode
PT
XX
XX Disclosure; Page 219; 227pp; English.
PS
XX
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AA206210) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 36 novel genes and their fragments (nucleic
CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 36
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AA206219 for described uses).
XX
XX Sequence 86 AA;
SQ
AAY38483 Length: 86 May 30, 2002 09:25 Type: P Check: 9545
..
1 MRQFYDPDE GLMKKGGLY FSDFWNKLDV GAILLFVAGL TCRLIPATLY
51 PGRVILSLDF ILFCLRLMHI FTISKTLGPK IIVKR
!!AA_SEQUENCE 1.0
ID AAY38484 standard; Protein; 27 AA.
XX
XX AAY38484;
AC
XX 18-OCT-1999 (first entry)
DT
XX
XX Human secreted protein encoded by gene No. 24.
DE
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
OS
XX
XX WO9935158-A1.
PN
XX 06-JAN-1999; 99WO-US00108.
PF
XX
XX 07-JAN-1998; 98US-0070704.
PR
XX 07-JAN-1998; 98US-0070657.
PR
XX 07-JAN-1998; 98US-0070658.
PR
XX 07-JAN-1998; 98US-0070692.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;
PI
XX Olsen HS, Rosen CA, Ruben SM, Soppet DR;
PI
XX
XX WPI; 1999-444190/37.
DR
XX N-PSDB; AA206242.
DR
XX
XX New isolated human genes and the secreted polypeptides they encode
PT
XX
XX Disclosure; Page 219; 227pp; English.
PS
XX
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AA206210) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 36 novel genes and their fragments (nucleic
CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 36
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AA206219 for described uses).
XX
XX Sequence 86 AA;
SQ
AAY38483 Length: 86 May 30, 2002 09:25 Type: P Check: 9545
..
1 MRQFYDPDE GLMKKGGLY FSDFWNKLDV GAILLFVAGL TCRLIPATLY
51 PGRVILSLDF ILFCLRLMHI FTISKTLGPK IIVKR
!!AA_SEQUENCE 1.0
ID AAY38484 standard; Protein; 27 AA.
XX
XX AAY38484;
AC
XX 18-OCT-1999 (first entry)
DT
XX
XX Human secreted protein encoded by gene No. 24.
DE
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
OS
XX
XX WO9935158-A1.
PN
XX 06-JAN-1999; 99WO-US00108.
PF
XX
XX 07-JAN-1998; 98US-0070704.
PR
XX 07-JAN-1998; 98US-0070657.
PR
XX 07-JAN-1998; 98US-0070658.
PR
XX 07-JAN-1998; 98US-0070692.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;
PI
XX Olsen HS, Rosen CA, Ruben SM, Soppet DR;
PI
XX
XX WPI; 1999-444190/37.
DR
XX N-PSDB; AA206242.
DR
XX
XX New isolated human genes and the secreted polypeptides they encode
PT
XX
XX Disclosure; Page 219; 227pp; English.
PS
XX
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AA206210) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 36 novel genes and their fragments (nucleic
CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 36
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AA206219 for described uses).
XX
XX Sequence 86 AA;
SQ
AAY38483 Length: 86 May 30, 2002 09:25 Type: P Check: 9545
..
1 MRQFYDPDE GLMKKGGLY FSDFWNKLDV GAILLFVAGL TCRLIPATLY
51 PGRVILSLDF ILFCLRLMHI FTISKTLGPK IIVKR
!!AA_SEQUENCE 1.0
ID AAY38484 standard; Protein; 27 AA.
XX
XX AAY38484;
AC
XX 18-OCT-1999 (first entry)
DT
XX
XX Human secreted protein encoded by gene No. 24.
DE
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
OS
XX
XX WO9935158-A1.
PN
XX 06-JAN-1999; 99WO-US00108.
PF
XX
XX 07-JAN-1998; 98US-0070704.
PR
XX 07-JAN-1998; 98US-0070657.
PR
XX 07-JAN-1998; 98US-0070658.
PR
XX 07-JAN-1998; 98US-0070692.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;
PI
XX Olsen HS, Rosen CA, Ruben SM, Soppet DR;
PI
XX
XX WPI; 1999-444190/37.
DR
XX N-PSDB; AA206242.
DR
XX
XX New isolated human genes and the secreted polypeptides they encode
PT
XX
XX Disclosure; Page 219; 227pp; English.
PS
XX
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AA206210) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 36 novel genes and their fragments (nucleic
CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 36
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AA206219 for described uses).
XX
XX Sequence 86 AA;
SQ
AAY38483 Length: 86 May 30, 2002 09:25 Type: P Check: 9545
..
1 MRQFYDPDE GLMKKGGLY FSDFWNKLDV GAILLFVAGL TCRLIPATLY
51 PGRVILSLDF ILFCLRLMHI FTISKTLGPK IIVKR
!!AA_SEQUENCE 1.0
ID AAY38484 standard; Protein; 27 AA.
XX
XX AAY38484;
AC
XX 18-OCT-1999 (first entry)
DT
XX
XX Human secreted protein encoded by gene No. 24.
DE
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
OS
XX
XX WO9935158-A1.
PN

```

PR 07-JAN-1998; 98US-0070658.
PR 07-JAN-1998; 98US-0070692.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;
XX
DR MPI; 1999-444190/37.
DR N-PSDB; AA206242.
XX
PT New isolated human genes and the secreted polypeptides they encode
XX
PS Disclosure; Page 219; 227pp; English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AA206210) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 36 novel genes and their fragments (nucleic
CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 36
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AA206219 for described uses).
XX
SQ Sequence 27 AA;
AAY38484 Length: 27 May 30, 2002 09:25 Type: P Check: 8317 ..
1 DECGLMKKG LYFSDFWNKL DVGAILL
!!AA_SEQUENCE 1.0
ID AAY36885 standard; Protein; 945 AA.
XX
AC AAY36885;
XX
DT 07-OCT-1999 (first entry)
XX
DE Protein involved in intermediate metabolism of polypeptides.
XX
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perinephritis;
KW nongonococcal urethritis; epidymitis; cervicitis; salpingitis;
KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX
OS Chlamydia trachomatis.
XX
PN WO9928475-A2.
XX
PD 10-JUN-1999.
XX
PF 27-NOV-1998; 98WO-IB01939.
XX
PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97ER-0015041.
PR 17-DEC-1997; 97ER-0016034.
XX
PA (GEST) GENSET.
XX
PI Griffals R;
XX
DR MPI; 1999-371125/31.
XX
PT Genome sequence of Chlamydia trachomatis
XX
PS Disclosure; Page 752-754; 1755pp; English.
XX
CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome

CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epidymitis, cervicitis, salpingitis,
CC perinephritis, Bartholinitis; pneumopathy in breast feeding infants;
CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.
XX
SQ Sequence 945 AA;
AAY36885 Length: 945 May 30, 2002 09:25 Type: P Check: 7481 ..
1 TRIALKEAVL LQRYELIRLI GKGGMGEVYL AHDKACSRRV ALKRIRLEDLS
51 GNALLRKRP RLAKIADLI HPGIVPYYSI CSDGEAVYYT MPYIEGFSLK
101 SLKSVWQKE VLSKELEERT SVKSFLLPIFD KICATVEYIH SKGYLHRDLK
151 PDNILLGLEF EVIVDWGAA IFKHAKELKL EQDDEAAVSF DERNICYSSM
201 TIPGKIYGP DYMAPESLLG VEASEKTDIY ALGLILYQL TLAFFYRRKK
251 GRKLSYRDVY LPIEMSPYR EIPPSLSQIA MKAIAINPAD RFSSIQELRLQ
301 ALQPYLOGP EWTVKATLMA KEKSCWKYYD PILSRFPV LASSPAQWYN
351 FMLSEVEISA STRVEYVYTK SAVHEGMGIL FLPSKEAERG EFYCGYGLWF
401 SVQNHLETVS LIKNGIEYQK KSQEMISQOY RFALLIEKSD NRIAVEVEQA
451 LFIHLIDYLP SLGNRLGVII QDLQMSNIA ISESIGALRV SCLAVPDAFL
501 SEKLYDQAAI FYRKIRDSFP GRKESYEAOE RLGVTLTLQI EEQGGDLTQA
551 LSSFYDLHGG AGAPLEYLCK ALVYQKNGSF VEEIRCLLFA LKRYSQHPEI
601 PRLEDHLCFR LYDSLHKHRS EALVFMILLI WIAPEKISVR EEKRPFLRIY
651 HKQATLFCQ VDKAPLOFRS SKMELFLSFV TAFSLFLEPL FRRAGELRDY
701 QALADIFYVA GVSNGREAFM QFSTALANVS DEITPPESLH NQKYAEIMPF
751 VKGVEALRNK DYQAKAKAFM GKTPTTLQLY ALDIRHIQAF LDEETIESFID
801 LLOAITDPAS EEBRDHILY IITQHLNMRD LERAYKLLND RPLDEELAE
851 YSEAFILMGC YLALTGDVVA VKAHFSRCRY KYGKSALIGK CVDGDIFDYL
901 DNLVWWEKKM TLFQSYFLLR CLNESPPRYE KYRQAVLSME NNFFD
!!AA_SEQUENCE 1.0
ID AAY38424 standard; Protein; 509 AA.
XX
AC AAY38424;
XX
DT 30-SEP-1999 (first entry)
XX
DE Human secreted protein.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.

PN WO9935158-A1.
XX
PD 15-JUL-1999.
XX
PE 06-JAN-1999; 99WO-US00108.
XX
PR 07-JAN-1998; 98US-0070704.
PR 07-JAN-1998; 98US-0070657.
PR 07-JAN-1998; 98US-0070658.
PR 07-JAN-1998; 98US-0070692.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Duan RD, Ebner R, Lafleur DW, NI J;
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;
XX
DR WPI; 1999-444190/37.
DR N-PSDB; AA206257.
XX
PT New isolated human genes and the secreted polypeptides they encode
XX
PS Claim 11; Page 198-199; 227pp; English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AA206210) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 36 novel genes and their fragments (nucleic
CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 36
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AA206219 for described uses).
XX
SQ Sequence 509 AA;

AAY38424 Length: 509 May 30, 2002 09:25 Type: P Check: 463 ..

1 MIAFPLLLTG LISFREKRLQ DVGTPARAR AFTAPVVF HNLISYFAF
51 LCLFAYVLMV DFQVPVSWCE CAIYIWLFSL VCEEMRQLFY DEDECGLMKK
101 AALYFSDFWN KLDVGAILLF VAGLTCRLIP ATLPGRAVIL SLDFILFCLR
151 LMHFTISKI LGPKIIIVKR MMKDVEFFLF LLAIVVVSFG VAKQAILIHN
201 ERRVDWLFEG AVYHSYLTIF GQIPGYIDGV NENPEHCSPN GTDPYKPKCP
251 ESDATQORPA FPEWLTVLLL CLYLFTNLL LNLLIAMFN YTFQOVQEH
301 DQYKFKQRHD LXEEYHGRPA APPPFILSH LQLFIKRVV KTPAKRHKQL
351 KNLEKNEEA ALISWEIYLL ENYFQNRQFO OKORPEQKIE DISNKVDAMV
401 DLIDLDPKLR SGSMEORLAS LEEQVAQTAR ALHWIVRTL ASGFSSEADV
451 PTLASOKAAE EPDAEPGGRK KTEEPGDSYH VNAHLLYPN CPVTRFPCPT
501 RRCPGRRSS

11AA_SEQUENCE 1.0
ID AAY34694 standard; Protein; 945 AA.

AC AAY34694;
XX
DT 13-SEP-1999 (first entry)
XX
DE C. pneumoniae protein involved in metabolism of polypeptides.
XX

KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
OS Chlamydia pneumoniae.
XX
PN WO9927105-A2.
XX
PD 03-JUN-1999.
XX
PE 20-NOV-1998; 98WO-IB01890.
XX
PR 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97FR-0014673.
XX
PA (GEST) GENSET.
XX
PI Griffals R;
XX
DR WPI; 1999-357842/30.
XX
PT Genome sequence of Chlamydia pneumoniae
XX
PS Page 692-694; Disclosure; 1912pp; English.
XX
CC AAY34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
SQ Sequence 945 AA;

AAY34694 Length: 945 May 30, 2002 09:25 Type: P Check: 6409 ..

1 ACIVCIDRED QRSLEBYDIV RIIGKGMGE VYLAYDPVCS RKVALKKIRE
51 DLAENPLKLR RFLREARIAA DLHPGVVPV YTIYSEKDPV YVTMPYIEGY
101 TLKTLKSWW OKESISKELA EKTSGVAFLS IFHKICCTIE YVHSGILHR
151 DLKPDNILLG LFSEAVILDW GAAVACGEEE DLDIDVSKE EYLSRMTIP
201 GRIVGPPDYM APERLGHPA SKSTDYALG VLYQMLTLS FPYRRKKGKK
251 IVLDGQRIPS PQEVAPYREI PPLLSAVVMR MLAVDPQERY SSVTELKEDI
301 ESHLKSPKW TLTALPPKK SSSWKLNPEI LSKYFPMLE VSPASWYSLA
351 ISNIESFSEM RLEYTISKKG LNEGFGILLP TSENALGGDF YQGYGFWLHI
401 KERTLSVSLV KNSLEIQRCS QDLES DKETP LIALEQHNHS LSLFVDGTTW
451 LIHMNYLPSR SGRVAIVRD MEDLEDIGI FESSGSLRVS CLAVPDAFLA
501 EKLIDRALVL YRRIASEFPG RKEGYEARFR AGITVLEKAS TDNNEQEFAL
551 AIEEFSKLHD GVAAPLEYLG KALVYQRLQE YNEETKSLLL ALKRYSQHPE
601 IFRLKDHYVY RLHESFYKRD RLALVEMILV LEIAPQAITP GQEEKILVWL
651 KDKSRATLFC LLDPYVLELR SSKMEFLSY WSGFIPHLS LFHRAWQSD
701 VRALIEIFYV ACDLHKWQFL SSCIDIFKES LEDQKATEEI VEFSEFDLGA
751 FLFAIQSIFN KEDAERFVS NDQLSPILLV YIFDLFANRA LLESQGEAIF

801 QALDIRSKV PENFYHDYLR NHEIRAHLMC RNEKALSTIF ENYTEKQIKD
851 EQHELFVLYG CYLALIQGAE AAKQHFVCR EDRIEPASLL ARNYNRLGLP
901 KDALSYQERR LLRQKFLYF HCLGNHDERD LCOTMYHLLT EEFQL
11AA_SEQUENCE 1.0
ID AAY05317 standard; Protein; 280 AA.
XX AAY05317;
AC
XX
DT 25-JUN-1999 (first entry)
XX
DE Human secreted protein bn97_1.
XX
KW Secreted protein; human; nutritional activity; cytokine; tissue growth;
KW cell proliferation; cell differentiation; immune stimulant; chemotaxis;
KW immune suppressant; haematopoiesis regulator; activin; inhibin; cadherin;
KW chemokinesis; haemostasis; thrombolysis; anti-inflammatory; gene therapy;
KW tumour invasion suppressor; tumour inhibitor.
XX
OS Homo sapiens.
XX
XX WO9913066-A1.
XX
XX PD 18-MAR-1999.
XX
XX PF 08-SEP-1998; 98WO-US18724.
XX
XX PR 08-SEP-1997; 97US-0929007.
XX
XX PA (GEMY) GENETICS INST INC.
XX
XX PI Agostino MJ, Evans C, Jacobs K, Lavallie ER, McCoy JM;
PI Merberg D, Racie LA, Spaulding V, Treacy M;
XX
XX DR WPI: 1999-229235/19.
XX
XX DR N-PSDB; AAX33810.
XX
XX PT New polynucleotides encoding secreted human proteins
XX
XX PS Claim 8; Page 79; 96pp; English.
XX
XX CC This sequence is a human secreted protein of the invention. The
CC secreted proteins were obtained from human adult placenta, foetal brain,
CC adult testes or adult blood cDNA libraries. The polynucleotides (Pns) and
CC proteins are predicted to have biological activities which would make
CC them suitable for treating, preventing or ameliorating medical conditions
CC in humans and animals, although no supporting data is given. Suggested
CC activities include nutritional activity, cytokine and cell
CC proliferation/differentiation activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity. The Pns are
CC also stated to be useful for gene therapy.
XX
XX SQ Sequence 280 AA;
AAY05317 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..
1 MQAKYSSTRD MLDDGDPTM SLHSQASATP RHEPRRTEH RAPSSTWRPV
51 ALTLTLTCLV LLIGLALGL LFFQYQLSN TGQDTISQME ERLGNTSQEL
101 QSLQVONIKL AGSLQHVAEK LCRELYNKAG AHRCSPTCEQ WKWHGDNQYQ
151 FYKDSKSWED CKYFCLSENS TMLKINKQED LEFAASQSYS EFFYSYWTGL
201 LRPDGKAWL WMDGTPFTSE LFHIIIDVTS PRSRDCAVIL NGMIFSKDCK
251 ELKRCVCERR AGMVKPESLH VPPETLGECD

11AA_SEQUENCE 1.0
ID AAW67815 standard; Protein; 57 AA.
XX
XX AC AAW67815;
XX
XX DT 25-MAR-1999 (first entry)
XX
XX DE Human secreted protein encoded by gene 9 clone HCEE50.
XX
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
XX FH Key location/qualifiers
FH MISC-difference 10
FH FT MISC-difference /label- unknown
FT MISC-difference 28
FT FT MISC-difference /label- unknown
XX
XX PN WO9842738-A1.
XX
XX PD 01-OCT-1998.
XX
XX PE 19-MAR-1998; 98WO-US05311.
XX
XX PR 30-MAY-1997; 97US-0050937.
XX PR 21-MAR-1997; 97US-0041276.
XX PR 21-MAR-1997; 97US-0041277.
XX PR 21-MAR-1997; 97US-0041281.
XX PR 21-MAR-1997; 97US-0042344.
XX PR 30-MAY-1997; 97US-0048069.
XX PR 30-MAY-1997; 97US-0048094.
XX PR 30-MAY-1997; 97US-0048095.
XX PR 30-MAY-1997; 97US-0048096.
XX PR 30-MAY-1997; 97US-0048099.
XX PR 30-MAY-1997; 97US-0048131.
XX PR 30-MAY-1997; 97US-0048135.
XX PR 30-MAY-1997; 97US-0048154.
XX PR 30-MAY-1997; 97US-0048160.
XX PR 30-MAY-1997; 97US-0048186.
XX PR 30-MAY-1997; 97US-0048187.
XX PR 30-MAY-1997; 97US-0048188.
XX PR 30-MAY-1997; 97US-0048350.
XX PR 30-MAY-1997; 97US-0048351.
XX PR 30-MAY-1997; 97US-0048352.
XX PR 30-MAY-1997; 97US-0048355.
XX PR 05-AUG-1997; 97US-0054804.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Brewer LA, Duan R, Ebner R, Ferrie AM, Florence KA;
PI Greene JM, Hu JS, Lafleur DW, Moore PA, Ni J, Olsen HS;
PI Rosen CA, Ruben SM, Shi Y, Young P;
XX
XX DR WPI: 1999-070066/06.
XX DR N-PSDB; AAX00619.
XX
XX PT New isolated human genes and the secreted polypeptides they encode -
XX PT useful for diagnosis and treatment of e.g. cancers, neurological
XX PT disorders, immune diseases, inflammation or blood disorders
XX
XX PS Claim 11; Page 277; 385pp; English.
XX
XX CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate

CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AAX00602) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 87 novel genes and their fragments (nucleic
CC acid sequences: AAX00611-X00724; amino acid sequences AAW67807-W68004)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 87
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAX00611 for described uses).

XX Sequence 57 AA;

AAW67815 Length: 57 May 30, 2002 09:25 Type: P Check: 5177 ..

1 MLVYDLYLX PKLWALATPQ KNGKGARKGD GTPAQAFWDF WSHLISADPQ

51 TWERAAP

11AA_SEQUENCE 1.0

ID AAB56913 standard; Protein: 275 AA.

XX AAB56913;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1491.

KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnery; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.

XX Homo sapiens.

PN WO200055174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM;

DR WPI: 2000-587513/55.

DR N-PSDB; AAF16116.

PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -

XX Claim 11; Page 1927-1928; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to

CC AAB57303 represent sequences used in the exemplification of the present
CC invention.

XX Sequence 275 AA;

AAB56913 Length: 275 May 30, 2002 09:25 Type: P Check: 7508 ..

1 KPEKGVHLN SDLPOMHNL IPLCAPNSLS QLPITDTIRK DSKEKKKRRR

51 SKTLWGTYH GMTLXSVTEG ASARKTQTPA AQPVP RPVSQ ARPPNQKRG

101 SRTPPIIIPA ATSLITMLN AKDLLQDLKF VPSDEKKKGQ CORENETLIQ

151 RRRDQMOPGG TAISVTVPYR VVDQPLKIMP QDWDRVAVF VQGPAMQFKG

201 WPMLLPDGSP VDIFAKIKAF HLKXDEVRLD PNVQKMDVTY LELSYHKRHL

251 DRPVFLRFWE TLDRYMVKHK SHLRF

11AA_SEQUENCE 1.0

ID AAB58349 standard; Protein: 100 AA.

XX AAB58349;

DT 14-MAR-2001 (first entry)

DE Lung cancer associated polypeptide sequence SEQ ID 687.

KW Human; lung cancer associated protein; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active; vulnery;
KW gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease.

XX Homo sapiens.

PN WO200055180-A2.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05918.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.

PI Ruben SM;

DR WPI: 2000-587514/55.

DR N-PSDB; AAF18225.

PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -

XX Claim 11; Page 1192; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnery; gastrointestinal
CC general; nephrotropic; antiinfective; gynaecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and

CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.
XX
SQ Sequence 100 AA;

AAB58349 Length: 100 May 30, 2002 09:25 Type: P Check: 8661 ..

1 SYNTLIPYC QCLFAAPHF FYIXTVLIF FCHWDCISDT LHXSLLAITW
51 KGSKGSGGA XRPGWXILQ NRNKTPOSIP LMPISQLFCC ISCLLFKKLP

!!AA_SEQUENCE 1.0
ID AAB03957 standard; Protein; 531 AA.

AC AAB03957;
DT 26-FEB-2001 (first entry)

DE Consensus human mesenchymal stem cell polypeptide.

KW Mesenchymal stem cell; cDNA; detection; genome; marker; locating;
KW genetic marker; linkage analysis; genetic disease; chromosome marker;
KW bone; gene therapy; genetic fingerprinting; peptide; immunogen;
KW antibody; human.

OS Homo sapiens.

PN WO200059933-A2.

PD 12-OCT-2000.

PF 31-MAR-2000; 2000WO-US08751.

PR 01-APR-1999; 99US-0127418.
PR 13-AUG-1999; 99US-0148800.

PA (OSIR-) OSIRIS THERAPEUTICS INC.

PI Van Den Bos C, Mbalaviele G;

DR WPI: 2000-664986/64.
DR N-PSDB; AAA54295, AAA54296.

PT Novel human mesenchymal stem cell polynucleotides and polypeptides
PT useful as reagents for chromosomal mapping, DNA fingerprinting and for
PT detecting presence of genetic mutations in diseases affecting bone
PT growth

PS Claim 16; Page 72-73; 73pp; English.

CC New human mesenchymal stem cell cDNA's and their fragments are useful
CC for detecting genes within a human genome i.e., for detecting the
CC presence of specific mRNA in a particular cell type and in genetic
CC linkage analysis, and for locating gene regions associated with
CC genetic disease, to control gene expression. The cDNA's, fragments
CC and expression products are useful as chromosome markers for
CC determining the location of genes within the genome and alleles
CC expressed, during the development of differentiated mesenchymal cell.
CC The cDNA's are also useful in detecting the presence of mutations
CC in genes leading to diseases affecting bone growth, by comparing the
CC sequences with those found in similar chromosomal locations in
CC mammals, especially humans afflicted with diseases affecting bone
CC growth, in human fingerprinting and in gene therapy. The polypeptides
CC encoded by the mesenchymal stem cell cDNA's or peptide fragments,
CC derivatives, analogs or cells expressing them are useful as
CC immunogens for the production of antibodies.

XX Sequence 531 AA;

AAB03957 Length: 531 May 30, 2002 09:25 Type: P Check: 4914 ..

1 MADVLSVLRO YNIOKKEIVV KGDEYIFGEF SWPKNVKTNV VVWGCTGKEGO

!!AA_SEQUENCE 1.0
ID AAB36865 standard; Protein; 1503 AA.

AC AAB36865;

DT 21-FEB-2001 (first entry)

DE Human putative Ca2+ channel protein Trpc7 protein.

KW mutTCC-1; sugar-phosphate hydrolase; NUDT9.

OS Homo sapiens.

PN WO200065056-A2.

PD 02-NOV-2000.

PF 26-APR-2000; 2000WO-US11319.

PR 26-APR-1999; 99US-0131051.

PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

PI Scharenberg AM;

DR WPI: 2000-687347/67.
DR N-PSDB; AAC68399.

PT Novel nucleic acids encoding a Mutr domain-containing polypeptide
PT present in human calcium channel protein Trpc7, useful for diagnostic
PT and therapeutic purposes -

PS Claim 39; Page 65; 69pp; English.

CC The present invention relates to human mutTCC-1. Therapeutics
CC involving mutTCC-1 may be useful for diagnosing and treating
CC conditions associated with aberrant levels of expression of (ii) and
CC for identifying agents that are useful for treating diseases associated
CC with pyrophosphohydrolase and/or sugar-phosphate hydrolase activity.

XX Sequence 1503 AA;

AAB36865 Length: 1503 May 30, 2002 09:25 Type: P Check: 5055 ..

1 MEPSALRRAG SEQEEGFEGL PRRVTDIGMV SNLRSSNSL FKSWRLOQCPF
51 GNNDKQESLS SWIPENIKKK ECVYFVSSSK LSDAGKVCQ CGYTHEQHL E
101 EATKPHTFQG TOWDPKRHHVQ EMPDPAFGDI VFTGLSQKVK KYVAVSQDTP
151 SSVIYHLMTQ HWGLDVPNLL ISVTGGAKNF NMKPRLSIF RRGIVKVAQT

201 TGAMITIGS HTGVAKQVGE AVRDESLSSS YKEGELITIG VATWGTVHRH
251 EGLIHTGSF PAEYILDEDEG QGNLTCLDSN HSHFILVDG THQGYVEIP
301 LRTRELEKFI EQTKERGVA IKIPICVVL EGGPGTLHTI DNATNGTPC
351 VVVEGSGRVA DVIAQVANLP VSDITISLIQ QKLSVEFQEM FETETESRIY
401 EWTKKIQDIV RRQOLLTVFR EKGDCQQDDV VAILQALLKA SRSQDHFGHE
451 NMDHQLKLAIV AMNRVDIARS EIPMDEWQWK PSDLHPTMTA ALISNKPEFV
501 KLEFENGVOI KEFVTWDTLL YLYENLDPSG LEHSKLQKVL VEDPERPACA
551 PAAPRLQMHV VAQVLRLLG DFTQPLYP RP RNDRLRLLL PVPVVKLVNQ
601 GVSLSRLYKR SSGHVTFTMD PIRDLIIWAI VQNRRELGI IMAQSQDCIA
651 AALACSKILK ELSKEEDTD SSEMLALAE EYEHRAIGVF TECYRKDEER
701 AQKLLTRVSE AMGKTTCLQL ALEAKDMKFV SHGIGQAFLT KVMWGQLSVD
751 NGLMRVTLGM LAFPLLLTGL ISFREKRLQD VGIPARARA FETAIVVFH
801 LNTISYFAFL CLEFAYVLMVD EQVPSPWCCE AIYLMLESLV CEMRQLFYD
851 PDEGLMKKA ALYFSDFWNK LDVGAILLV AGLTCLRLPA TLVPGRVILS
901 LDFILFCLRL MHITITSKTG GPRIIIVKRM MKDVEFFFLF LAVWVVSFGV
951 AKQALIHNE RRVDMLEFGA VYHSYLTIFG QIPGYIDGVN FNEHCSPNG
1001 TDPKPKCPE SDATQQRPAF PEMLTVLLIC LYLLFTNILL LNLIIAMFN
1051 TFOQVQEHND QIMKFRQHRDL IEEYHGRPA PPFPILLSHL QLEIKRVVLK
1101 TPARKHKQLK NKEKNEEA LLSWEIYLYKE NYLQNRQFOQ KQRPQKIED
1151 ISNKVDAMVD LLDLDPLKRS GSMQORLASL EEOVAQTARA LHWIVRTLRA
1201 SGFSEADVP TLASOKAAEE PDAEPGGRRK TEPEGDSYHV NARHLLYPNC
1251 PVTRPPVPNE KVPETEFLI YDPPFYTAER KDAAMDPMG DTLEPLSTIQ
1301 YNVVDGLRDR RSFHPYTVQ AGLPINPMGR TGLRGGSLS CFGPNHTLYP
1351 MVTRRARNED GAICRKSIAK MLEVIVVKLP LSEHWALPGG SREPGEMLP
1401 KLRRLRQEH WPSRENLLKC GMEVYKGYMD DPRINTNAMI ETVAVSVHFQ
1451 DQNDVELNRL NSNLHACDSG ASIRWQVDR RIPLEYANHKT LIQKAAAEFG
1501 AHY

11AA_SEQUENCE 1.0
ID AAB12123 standard; Protein; 344 AA.
XX
AC AAB12123;
XX
DT 02-FEB-2001 (first entry)
XX
DE Hydrophobic domain protein from clone HP10608 isolated from Saos-2 cells.
XX
KW Human; secreted protein; membrane protein; hydrophobic domain;
KW proliferation control; differentiation induction; material transport;
KW biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200029448-A2.

XX
PD 25-MAY-2000.
XX
PF 17-NOV-1999; 99WO-JP06412.
XX
PR 17-NOV-1998; 98JP-0326255.
PR 22-DEC-1998; 98JP-0364315.
PR 16-MAR-1999; 99JP-0069811.
PR 27-APR-1999; 99JP-0119299.
PR 19-MAY-1999; 99JP-0138169.
XX
PA (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX
PI Kato S, Kimura T;
XX
DR WPI; 2000-387753/33.
DR N-PSDB; AAA60185, AAA60195.
XX
XX
PT Proteins comprising hydrophobic regions, such as secretory and membrane
PT proteins, useful in research and diagnostics and having various
PT activities e.g. immunomodulatory, antiinflammatory, chemokinetic,
PT hemostatic, thrombolytic -
XX
PS Claim 1; Page 188-189; 410pp; English.
XX
CC Secretory proteins play important roles in the proliferation control, the
CC differentiation induction, the material transport and the biophylaxis of
CC cells. Membrane proteins have important roles as signal receptors, ion
CC channels and transporters. The present sequence is a human protein which
CC has at least one hydrophobic domain. This protein may be a secretory or a
CC membrane protein. The present protein may have cytokine and cell
CC proliferation/differentiation activity, immune stimulating or suppressing
CC activity, haematopoiesis activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, anti-inflammatory activity and tumour
CC inhibition activity. The present protein could therefore be used for
CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's
CC disease, and cancer.
XX
SQ Sequence 344 AA;
XX

AAB12123 Length: 344 May 30, 2002 09:25 Type: P Check: 1439 ..

1 MQPPPPPLG DCLRDWEDIQ QDFONIQETH RLYRKLEEL TKLQNNCTSS
51 ITRQKKRIQE LALALKCKP SLPAEAGAA QELEQMKER QGLEFDMEAY
101 LPKKNGLYLS LVLGNNVTLL LSKQAKFAYK DEYEKFKLYL TIILILISFT
151 CRFLINSRYT DAAFNELIWW YCYTLTIRES ILINNGSRIK GWWVFHHYVS
201 TFLSGVMLTW PDGLMYQKFR NQFLSFSMYQ SFVQPLQYYY QSGCLYLRRA
251 LGERHTMDLT VEGFQSKMMWR GLTFLLPFLF FGHFWQLFNA LTFENLAQDP
301 QCKEWOYIMC GPFLLLFLG NFFTILRVVH HKFHSORHGS KKDD

11AA_SEQUENCE 1.0
ID AAB41885 standard; Protein; 251 AA.
XX
AC AAB41885;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1649 polypeptide sequence SEQ ID NO:3298.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antiparinsonian; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW

KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.

OS Homo sapiens.
XX
XX WO200058473-A2.
PN
XX
XX 05-OCT-2000.
PD
XX
XX 31-MAR-2000; 2000WO-US08621.
PF

XX 31-MAR-1999; 99US-0127607.
PR
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shinkets RA, Leach M;
PI
XX
XX WPI: 2000-602362/57.
DR
DR N-PSDB; AAC76094.

XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX

PS Claim 11; Page 2491; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX

Sequence 251 AA;

AAB41885 Length: 251 May 30, 2002 09:25 Type: P Check: 2227 ..

1 GHLVCVLCVC MYLCVCVCLC VCALCLCVCL CECLMVSFCV CVCICVCVXL
51 CACMCLDVCF CMCLCVCLYV CICYVVCVCH FVCFMVCLSA CLCTIVSPCV
101 CLCVCICXCL CMCVRCVSV CVCVCIEREG ERKGATDGA WKVYPHSQPM
151 EESVNPPETGQ DQIMWCLADS GNVTFHURMG LHFLEKECRS WSLKECFEPP
201 FVIERAOPCV HMLTVTNLRY GDSHRETEG TADEQESGG TSLPLGPNPO
251 L

11AA_SEQUENCE 1.0
ID AAB42619 standard; Protein; 280 AA.

XX
XX AAB42619;
AC
XX 08-FEB-2001 (first entry)
DT
XX
XX Human ORFX ORF2383 polypeptide sequence SEQ ID NO:4766.
DE

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.

XX
XX Homo sapiens.
OS
XX
XX WO200058473-A2.
PN
XX
XX 05-OCT-2000.
PD
XX
XX 31-MAR-2000; 2000WO-US08621.
PF

XX 31-MAR-1999; 99US-0127607.
PR
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shinkets RA, Leach M;
PI
XX
XX WPI: 2000-602362/57.
DR
DR N-PSDB; AAC76828.

XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX

PS Claim 11; Page 3947-3948; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX

Sequence 280 AA;

AAB42619 Length: 280 May 30, 2002 09:25 Type: P Check: 2818 ..

1 LQAKYNSTRD MLDGDDITM SLHSQASATT RHEPRRTEH RAPSSTWBPV
51 ALTLTLCLV LLIGLALGL LFFQYYQLSN TGQDTISQME ERLGNTSQEL
101 QSLQVQNIKL AGSLQHVAEK LCRELYNKAG AHRCSPCTEQ WKWHGNCYQ
151 FYKDSKSWED CKYFCLSENS TMLKINKQED LEPAASQSYS EFFYSYWTGL
201 LRPDSGKAWL WMDGTPEPTESE LEHIIIDVYS PRSRDCAVAIL NGMIFSKDCK
251 ELKRCVCERR AGMVKPESLH VPPETLGEGD

11AA_SEQUENCE 1.0
ID AAB32818 standard; Protein: 113 AA.

XX AAB32818;

DT 25-JAN-2001 (first entry)

DE Eucalyptus grandis transcription factor protein sequence #276.

XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB.

XX Eucalyptus grandis.

PN WO200053724-A2.

PD 14-SEP-2000.

PF 09-MAR-2000; 2000WO-US06112.

PR 11-MAR-1999; 99US-0266513.

PR 18-AUG-1999; 99US-0149485.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

PI Wood M, McGrath A, Shenk MA, Glenn M;

DR WPI; 2000-579369/54.

PT New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide

PS Claim 8; Page 316; 747pp; English.

XX The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
CC transcription factor. The transcription factor may be used to produce a
CC plant having modified gene expression such as a woody plant e.g. a
CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
CC to modify the activity of a polypeptide in a plant. The transcription
CC factors of the present invention are members from the following families
CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
CC zipper, LIM domain, AP2 and EREBS, zinc finger domains of type 2
CC Cys2His2, CCAAT box elements and MYB.

XX Sequence 113 AA;

AAB32818 Length: 113 May 30, 2002 09:25 Type: P Check: 7107 ..

1 EIKNYWNTRI KRLQRTGMPI YPTEVCLQYS SENQETHNMG NLHTAGEDNC

51 DLSQADPLEI PEVDFRKLLEL HLGSSFWST LLDVPCGFG REAMCLSDAY
101 CLPEPSSRSP KRL

11AA_SEQUENCE 1.0
ID AAB38371 standard; Protein: 88 AA.

AC AAB38371;

DT 31-JAN-2001 (first entry)

DE Human secreted protein encoded by gene 51 clone HLDQC46.

XX Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
KW cytostatic; cardiant; vasotropic; cerebroprotective; neuroprotective;
KW nootropic; antibacterial; virucide; fungicide; ophthalmological; human;
KW vulnerary; gene therapy; infection; secreted protein.

XX Homo sapiens.

PN WO200061623-A1.

PD 19-OCT-2000.

PF 06-APR-2000; 2000WO-US08979.

PR 09-APR-1999; 99US-0128693.

PR 26-APR-1999; 99US-0130991.

PA (HUMA-) HUMAN GENOME SCI INC.

PA Ruben SM, Ni J, Komatsoulis GA, Rosen CA, Soppet DR, Shi Y;

PI Lafleur DM, Olsen HS, Ebner R, Florence KA, Moore PA, Birse CE;

PI Young PE;

DR WPI; 2000-647418/62.

XX New nucleic acid molecules encoding 62 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -

PS Claim 11; Page 636; 716pp; English.

XX Sequences AAB38321-B38396 represent the amino acid sequences of 62
CC human secreted proteins encoded by the genes AAC69512-C69587. The genes
CC and proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.
CC rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms
CC of the breast or liver; (c) cardiovascular disorders e.g. cardiac
CC arrest; (d) cerebrovascular disorders e.g. cerebral ischemia; (e)
CC angiogenesis; (f) nervous system disorders e.g. Alzheimer's disease; (g)
CC infections caused by bacteria, viruses and fungi; and (h) ocular
CC disorders e.g. corneal infection. The polypeptides can also be used to
CC aid wound healing and epithelial cell proliferation, to prevent skin
CC aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis.

XX Sequence 88 AA;

AAB38371 Length: 88 May 30, 2002 09:25 Type: P Check: 3048 ..

1 MDLVEGFQGS WMWRGLTFLL PFLFGHFWD LFNALTFNL AQDPCKEMQ

51 VLMCGFPPLL LFLGNFTTTL RVVHKKFHSQ RHGSKKDX

11AA_SEQUENCE 1.0

ID AAB17077 standard; Peptide: 12 AA.

AC AAB17077;

XX 31-OCT-2000 (first entry)
DT
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:133.
XX
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
XX WO200024782-A2.
PN
XX 04-MAY-2000.
PD
XX 25-OCT-1999; 99WO-US25044.
PF
XX 23-OCT-1998; 98US-0105371.
PR
XX 22-OCT-1999; 99US-0428082.
PT
XX (AMGE-) AMGEN INC.
PA
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
XX WPI; 2000-350702/30.
DR
XX
XX Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
XX
PS Claim 39; Page 241; 608pp; English.
XX
XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;
AAB17077 Length: 12 May 30, 2002 09:25 Type: P Check: 6124 ..
1 QETFSDYWKL LP
11AA_SEQUENCE 1.0
ID AAB17078 standard; Peptide; 12 AA.
XX
AC AAB17078;
XX
XX 31-OCT-2000 (first entry)
DT
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:134.
XX
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW

KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
XX WO200024782-A2.
PN
XX 04-MAY-2000.
PD
XX 25-OCT-1999; 99WO-US25044.
PF
XX 23-OCT-1998; 98US-0105371.
PR
XX 22-OCT-1999; 99US-0428082.
PT
XX (AMGE-) AMGEN INC.
PA
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
XX WPI; 2000-350702/30.
DR
XX
XX Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
XX
PS Claim 39; Page 241; 608pp; English.
XX
XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;
AAB17078 Length: 12 May 30, 2002 09:25 Type: P Check: 6146 ..
1 QPTFSDYWKL LP
11AA_SEQUENCE 1.0
ID AAB17079 standard; Peptide; 12 AA.
XX
AC AAB17079;
XX
XX 31-OCT-2000 (first entry)
DT
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:135.
XX
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
XX WO200024782-A2.
PN

XX 04-MAY-2000.
PD
XX
XX 25-OCT-1999; 99WO-US25044.
PF
XX
XX 23-OCT-1998; 98US-0105371.
PR
XX 22-OCT-1999; 99US-0428082.
PT
XX
XX (AMGE-) AMGEN INC.
PI
XX Feige U, Liu C, Cheetham J, Boone TC;
DR
XX WPI; 2000-350702/30.
PT
XX Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
XX
PS Claim 39; Page 241; 608pp; English.
XX
XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;
AAB17079 Length: 12 May 30, 2002 09:25 Type: P Check: 5978 ..
1 MPREFDYWEG LN
11AA_SEQUENCE 1.0
ID AAB17080 standard; peptide: 12 AA.
XX
AC AAB17080;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:136.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
XX 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX

PA (AMGE-) AMGEN INC.
XX
XX Feige U, Liu C, Cheetham J, Boone TC;
XX
XX WPI; 2000-350702/30.
DR
XX
XX Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
XX
PS Claim 39; Page 242; 608pp; English.
XX
XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;
AAB17080 Length: 12 May 30, 2002 09:25 Type: P Check: 6151 ..
1 VQNFIDYWTQ QF
11AA_SEQUENCE 1.0
ID AAB17081 standard; peptide: 12 AA.
XX
AC AAB17081;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:137.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
XX 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
XX (AMGE-) AMGEN INC.
PI
XX Feige U, Liu C, Cheetham J, Boone TC;
DR
XX WPI; 2000-350702/30.
XX
XX Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and

PT autoimmune diseases -
XX
PS Claim 39; Page 242; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytosstatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.

XX Sequence 12 AA;

AAB17081 Length: 12 May 30, 2002 09:25 Type: P Check: 5993 ..

1 TGPATHTYWA TF

11AA_SEQUENCE 1.0
ID AAB17082 standard; Peptide; 15 AA.

XX AAB17082;
AC
XX

DT 31-OCT-2000 (first entry)

XX Mdm/hdm antagonist peptide sequence SEQ ID NO:138.

DE
XX
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.

OS Synthetic.

XX WO200024782-A2.

XX 04-MAY-2000.

PD 25-OCT-1999; 99WO-US25044.

XX 23-OCT-1998; 98US-0105371.

PR 22-OCT-1999; 99US-0428082.

XX (AMGE-) AMGEN INC.

PA Feige U, Liu C, Cheetham J, Boone TC;

PI WPI; 2000-350702/30.

DR
XX
XX Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -

XX Claim 39; Page 242; 608pp; English.

XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,

CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytosstatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.

XX Sequence 15 AA;

AAB17082 Length: 15 May 30, 2002 09:25 Type: P Check: 9093 ..

1 IDRAPTFRDH WFALV

11AA_SEQUENCE 1.0
ID AAB17083 standard; Peptide; 15 AA.

XX AAB17083;
AC
XX

DT 31-OCT-2000 (first entry)

XX Mdm/hdm antagonist peptide sequence SEQ ID NO:139.

DE
XX
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.

OS Synthetic.

XX WO200024782-A2.

XX 04-MAY-2000.

PD 25-OCT-1999; 99WO-US25044.

XX 23-OCT-1998; 98US-0105371.

PR 22-OCT-1999; 99US-0428082.

XX (AMGE-) AMGEN INC.

PA Feige U, Liu C, Cheetham J, Boone TC;

PI WPI; 2000-350702/30.

DR
XX
XX Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -

XX Claim 39; Page 243; 608pp; English.

XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytosstatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are

CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 15 AA;

AAB17083 Length: 15 May 30, 2002 09:25 Type: P Check: 9428 ..
1 PRPALVFADY WETLY
11AA_SEQUENCE 1.0
ID AAB17084 standard; Peptide; 15 AA.
AC AAB17084;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:140.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 243; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 15 AA;

AAB17084 Length: 15 May 30, 2002 09:25 Type: P Check: 8833 ..
1 PAFSRFWSKL SAGAH
11AA_SEQUENCE 1.0
ID AAB17085 standard; Peptide; 15 AA.
AC AAB17085;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:141.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 243; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 15 AA;

AAB17085 Length: 15 May 30, 2002 09:25 Type: P Check: 8896 ..
1 PAFSRFWSKL SAGAH
11AA_SEQUENCE 1.0
ID AAB17086 standard; Peptide; 10 AA.
XX

AC AAB17086;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:142.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
PI WPI; 2000-350702/30.
XX
DR
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
XX
PS Claim 39; Page 243; 608pp; English.
XX
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 10 AA;
AAB17086 Length: 10 May 30, 2002 09:25 Type: P Check: 4557 ..
1 PFXFDYWXXL
11AA_SEQUENCE 1.0
ID AAB17089 standard; Peptide; 12 AA.
XX
AC AAB17089;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:145.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;

KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
PI WPI; 2000-350702/30.
XX
DR
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
XX
PS Claim 39; Page 244; 608pp; English.
XX
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;
AAB17089 Length: 12 May 30, 2002 09:25 Type: P Check: 6124 ..
1 QETFSDYWKL LP
11AA_SEQUENCE 1.0
ID AAB17090 standard; Peptide; 12 AA.
XX
AC AAB17090;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:146.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.

PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US25044.
 XX
 PR 23-OCT-1998; 98US-0105371.
 PR 22-OCT-1999; 99US-0428082.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheetham J, Boone TC;
 XX
 DR WPI; 2000-350702/30.
 XX
 PT Novel composition of matter comprising an Fc domain and
 PT pharmacologically active peptides, useful for treating cancer and
 PT autoimmune diseases -
 XX
 PS Claim 39; Page 245; 608pp; English.
 XX
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
 CC where P1, P2, P3, and P4 = are each independently sequences of
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
 CC independently linkers; and a, b, c, d, e, and f = are each independently
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
 CC have cytosolic, antitumoral, thrombolytic and immunosuppressive
 CC activities. DNAs, vectors and host cells from the present invention can
 CC be used for producing pharmaceutical compositions. The compositions are
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer
 CC half-life or incorporate functions such as Fc receptor binding, protein
 CC A binding, complement fixation, and possibly placental transfer. AAA69443
 CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
 CC sequences used in the exemplification of the present invention.
 XX
 SO Sequence 12 AA;
 AAB17090 Length: 12 May 30, 2002 09:25 Type: P Check: 6146 ..
 1 QPTFSQYWK L P
 11AA_SEQUENCE 1.0
 ID AAY92944 standard; Protein; 1503 AA.
 XX
 AC AAY92944;
 XX
 DT 08-NOV-2000 (first entry)
 XX
 DE Human TRPC7 protein.
 XX
 KW Transmembrane protein; TRPC7; brain; transient receptor potential; TRP;
 KW calcium channel function; human; gene therapy; periodic psychosis;
 KW mutation.
 XX
 OS Homo sapiens.
 XX
 PN WO200029571-A1.
 XX
 PD 25-MAY-2000.
 XX
 PE 11-NOV-1999; 99WO-JP06289.
 XX
 PR 12-NOV-1998; 98JP-0321200.
 XX
 PA (EIKE) EIKEN KAGAKU KK.
 XX
 PI Shimizu N, Nagamine K;
 XX

DR WPI; 2000-387784/33.
 DR N-PSDB; AAA11284.
 XX
 PT Nucleic acids encoding transmembrane protein TRPC7 expressed in brain
 PT and homologous to transient receptor potential protein useful in the
 PT treatment of associated diseases such as periodic psychosis -
 XX
 PS Claim 1; Page 64-71; 77pp; Japanese.
 XX
 CC The invention relates to the isolation of a nucleic acid encoding a
 CC transmembrane protein TRPC7 which is expressed in brain and is homologous
 CC to transient receptor potential (TRP) protein. This suggests that the
 CC TRPC7 protein may have a calcium channel function. This sequence
 CC represents the human TRPC7 protein. The DNA and protein can be used
 CC in the diagnosis and treatment of disorders associated with TRPC7,
 CC especially the screening, monitoring and treatment (by gene therapy)
 CC of periodic psychosis, which appears to be associated with mutations
 CC in the TRPC7 gene.
 XX
 SQ Sequence 1503 AA;
 AAY92944 Length: 1503 May 30, 2002 09:25 Type: P Check: 5055 ..

1 MEPSALRKAG SEQEGFEGF PRRVTDLGMV SNLRSSNSL FKSRLQCPF
 51 GNNDKQESLS SWIPENIKKK ECVYFVSSK LSDAGKVVCQ CGYTHEQHLE
 101 EATKPHTFQG TQWDPKHHVQ EMPDADFQDI VFTGLSQKVK KYVRVSQDTP
 151 SSVIYHMTQG HWGLDVPNLL ISVTGAKNF NMKPLKSIK RRLGVKVAQT
 201 TGAWIITGGS HTGVMKOVGE AVRDFSLSS YKEGELITIG VATWGTVHRR
 251 EGLIHPTGSF PAEYILDEDS QGNLTCLDSN HSHFILVDDG THGOYGEIIP
 301 LTRLEKFTIS EQTKERGVA IKIPVCVYL EGPGTLHTI DNATTNGTPC
 351 VVEGSGRVA DVIAQVANLP VSDITISLQ OKLVSFFQEM FETTESRIY
 401 EWTKKIQDIV RRRQLTVFR EGKDGQDQVD VALQALKA SRSQDHFGHE
 451 NMDHQLKLAIV AMNRVDIARS EIFMDEWQW PSDLHPTMTA ALISNKPEFV
 501 KLFLENGVQL KEFVTWDTLL YLYENLDPSK LFHSLKQKVL VEDPERPACA
 551 PAAPRLQMHV VAQVRELGL DFTQPLPRP RHNDRLRLLL PVPHVKLVNQ
 601 GVSLSRLYKR SSGHVTFTMD PIRDLLIMAI VQNRRELAGI IWAQSQDCLA
 651 AALACSKILK ELSKEEDTD SSEMLALAE EYEHRAGVF TECYRKDEER
 701 AQKLLTRVSE AMGKTTCLQL ALEAKDMKEV SHGGIQAFLT KVMWGQLSYD
 751 NGLMRTVLCM LAFPLLLTGL ISFREKRLQD VGTAPARARA FETAPVVVFH
 801 LNLISYFAFL CLFAYVLMVD FQPVPSWCEC AIYIMLFSLV CEEMRQLFTD
 851 PDECGLMKKA ALYFSEFWNK LDVGAILLFV AGLTCRLIPA TLYPGRVILS
 901 LDFILFCLRL MHIFTISKTL GPKIIVKRM MKDVFFFLFL LAVWVVSFGV
 951 AKQAILIHNE RRYDWLFRGA VYHSYLTIFG QIPGYIDGVN FNEPHCSPPG
 1001 TDPYKPKCP ESDATQORPAF PEMLVLLLC LYLFTNILL LNLIIAMFNY
 1051 TFOQVOEHTD QIWKQFQRHDL IEEYHGRPA PPPIILLSH QLFIKRVVLK
 1101 TPAKRHKQLK NLEKNEEAA LLSWEIYLKE NYLQNRQFQ KQREPKIED
 1151 ISNKVDAMVD LLDLDPLKRS GSMQRLASL EGVQAQTARA LHMIVRTLRA
 1201 SGFSSEADV PTLASQKAAE PDAEPGRKK TEEPGDSYHV NARHLLYPNC

1251 PVTREPVPE KVPWETEFLL YDPFFYTAER KDAAMDPMG DTLEPLSTIQ
1301 YNVVDGLDR RSFHGPIYQ AGLPLNPMGR TGLRGGSLS CEGPNHTLYP
1351 MYTRWRNED GAICRKSIRK MLEVLYVKLP LSEHWALPGG SREPGEMLPR
1401 KIKRILRQEH WPSFENLKC GMEVYKGYMD DPRNTDNAMI ETVAVSVHFQ
1451 DQNDVELNRL NSNLHACDSG ASIRWQVYDR RIPLYANHKT LIQKAAAEFG
1501 AHY
11AA_SEQUENCE 1.0
ID AAG17202 standard; Protein: 584 AA.
XX
XX AAG17202;
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 18134.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.

PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 08-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.

PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

XX (CERE-) CERES INC.

XX Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;

XX WPI; 2000-507395/46.
DR N-PSDB; AAC37642.

XX New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -

PS Claim 19; SEQ ID 18134; 344pp + CD-ROM; English.

XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.

XX Sequence 584 AA;

AAG17202 Length: 584 May 30, 2002 09:25 Type: P Check: 7959 ..

1 YFEIIVLFL LIRIRNSSF KLFINLVFEF WSEIYMGSGS VTSIAPGFRF
51 HPTDELVRY YLKRKICNKP FKFDALSVTD VYKSEPMWLP DKSRLKSRLD
101 EMYFFSMLNK KYRNGSKTNR ATEMGYWKTT GKDRREILNGS KVGGMKRTLV
151 YHKGRAPRGE RTNWVMHEYR LVDQDLDTKG VHQDAFVLCR IFQKSGSGPK
201 NGEQYGAPFV EEWEEEDDM TFPDQEDLG SEDHYVYHMD DIDQKSENV
251 VYDAIPIPLN FIHGESSNNV ETNYSDSINY IQOTGNVMDS GGYFQPPAES
301 YEKDQKPIIR HRDGLQNEG ICGGVQDKHS EILQSSDNIF GTDTSQYDNF
351 PVESNYLIGE AFLDPNSNLT ENDGLYLETN DLSSTQODGF DFEDYLFEPD

401 EFPDPSOLMG NEDVFPDQEE LFQEVETKEL EKEETSRSKH VVEKEKDEA
451 SCSKQVDADA TEEFPDYKYP ILKKASHMLG AIPAPLANAS EEP TKDAIR
501 LHAAQSSGSV HTYAGMITIS DSNMGWSYK NENIDLILSL GLVOGNTAPE
551 KSGNSSAWAM LIFMCFWVLL LSVSEKVSIL VSSR

11AA SEQUENCE 1.0
ID AAG17203 standard; Protein; 549 AA.

XX AAG17203;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 18135.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 04-MAY-1999; 99US-0132048.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132486.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135124.

PR 24-MAY-1999; 99US-0135353.

PR 25-MAY-1999; 99US-0135629.

PR 27-MAY-1999; 99US-0136021.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137502.

PR 04-JUN-1999; 99US-0137724.

PR 07-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 14-JUN-1999; 99US-0138847.

PR 16-JUN-1999; 99US-0139119.

PR 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.

PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.

PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

(CERE-) CERES INC.

Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
Zheng L, Dumas J;

WPI; 2000-507395/46.
N-PSDB; AAC37642.

New sequence determined DNA fragments (SDFs) from different plant
species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
protein coding sequences, untranslated regions, or as 3' termination
sequences -

Claim 19; SEQ ID 18135; 344pp + CD-ROM; English.

The present sequence is a putative protein fragment from
Arabidopsis thaliana. Its coding sequence was isolated by carrying out
RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
library which could then be sequenced, allowing the putative protein
sequence(s) to be obtained. This sequence may be useful for protein
identification and for aiding in the elucidation of signal transduction
and metabolic pathways. Its coding sequence has a use in the control of
gene expression as a promoter, coding sequence, 3'UTR or termination
sequence, for controlling the behaviour of a gene within the chromosome,
as a tool for use in genetic mapping, including a use in hybridisation
assays, for recognition or isolation of similar DNA fragments, or for
the identification of a particular organism.

Sequence 549 AA;

AAG17203 length: 549 May 30, 2002 09:25 Type: P Check: 1255 ..

1 MGRGSVTSIA PGRFHPTDE ELVRYTLKRR ICNKPFRDA ISVTDVYKSE
51 PWDLPDKSRL KSRDLEWYFF SMLNKKYRNG SKTNRATEMG YWKTGKDRRE
101 IINGSKVVG M KTLVYHKGR APRGERTNMV MHEYRLVDQD LDKTGVAQDA
151 FVLCRIFOKS GSGPKNGEYV GAPFVEEWE EEDDMTFYVD QEDIGSEDHV
201 YVHMDIDQK SENFVYDAI PIPLNFHGE SSNNVEFNYS DSINYIQGTG
251 NYMDSGGYFE QPAESYEKQ KPIIRHRDGS LQNEGIGCV QDKHSELIQS
301 SDNIFGTDTS CYNDPVEESN YLIGEAFLDP NSNLLENDCL YLETNDLSST
351 QODGFDFEDY LTFEDEFDP SOLMGNEDEV FDQDELPOEV ETKELEKEET
401 SRSKHVEEK EKDEASCSKQ VDADATEFEP DYKYPILKKA SHMLGAIAPAP
451 LANASEFPK DAIRLHAAQ SSGSVHYTAG MITISDSNMG WSYGKNENLD

501 LILSLGLVQG NTAPKSGNS SAWAMLIIFMC FWLLLSVSF KVSILVSSR
IIAA_SEQUENCE 1.0
ID AAG17204 standard; Protein; 478 AA.
XX
AC AAG17204;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18136.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.

PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171107.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 26-JAN-2000; 2000US-0176910.
PR 27-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.

PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX
PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
DR WPI; 2000-507395/46.
DR N-PSDB; AAC37642.
XX
PT New sequence determined DNA fragments (SDfs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
PS Claim 19; SEQ ID 18136; 344pp + CD-ROM; English.
XX
CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
CC
XX
SQ Sequence 478 AA;
AA017204 Length: 478 May 30, 2002 09:25 Type: P Check: 530 ..
1 MLNKKYRNGS KTNRATEMGY WKTGKDREI LNSKVVGMK KTLVYHKGRA
51 PRGERTNMVM HEYRLVDQDL DKTGVHDAF VLCRIFOKSG SGPKNGEQYG
101 APFVEEEMEE EDDMTFVPDQ EDLGSSEHNY VHMDIDQKS ENFVVDATIP
151 IPLNFIHGES SNNVETNYSY SINYIQGTGN YMDSGGYFEQ PAESYEKQDK
201 PIRHRDGL QNEGIGGVQ DKHSEILQSS DNIFGDTSC YNDFPVESNY
251 LIGFAFLDPN SNLLENDGLY LETNDLSTQ QDGFDFEDYL TFFDETDPDS
301 QLMGNEDVFF DQELFOEVE TKLEKEETS RSKHVEEKE KDEASCSKQY
351 DADATEFEPD KYRPLKKAS HMLGAIPAPL ANASEFPTKD AAIRLHAQOS
401 SGSVHTAGM ITISDSNMGW SYGKNENLDL ILSLGLVQGN TAPEKSGNS
451 AWAMLIFMCF WYLLLSVSFK VSILVSSR
11AA_SEQUENCE 1.0
ID AAG21029 standard; Protein; 504 AA.
XX
AC AAG21029;
XX
DT 17-OCT-2000 (first entry)

XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23437.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159658.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.

PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 03-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.

PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX
PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
DR WPI; 2000-507395/46.
DR N-PSDB; AAC39118.
XX
XX
PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
PS Claim 19; SEQ ID 23437; 344pp + CD-ROM; English.
XX
CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 504 AA;
AAG21029 Length: 504 May 30, 2002 09:25 Type: P Check: 5419 ..
1 MTFAGETHV RRRDLQNL SLFQSLFW SWFFSPLSS LSQNPKLSFP
51 LVFFSEGLIC DGRVNYSTS PAPHNSDAD PLNPTRRES HRSIILSKF
101 LEQFFIDTSF GSRSAVLAVT PLVPSHRISL MARDSLGPSV VPPVRTSTT
151 HDELHCQVF FPGASPTT SDLYDGCALP SLSTVTTLG FCGMGLNFLD
201 GPVATIHAFG GLGDSQFGGP KTLNLIWPIIS RFALLRVLST LSLNHCFTIT
251 DGRGLHSIVG TLVLEDPCHN NTRNLLHRVD WLPTDVSLVL KLSGLVLISN
301 HRLCDAAHAF VFICDHIMSL NLPTIMLTHP VPCRLLDAS QPSWNCWLKL
351 LRPALSTLFS SPMNPYYLCH LIRRLSVLWT CNLCRLHYV YSSKEIPLSK
401 PYSSKEISSL KLEPGEKSHT AASSLEHFNK SEYASIRAY YDHKLIRDFS
451 KLVFKLESLLQ VPIDPLSLAT FSKLCLEDS WNLVLYLSCN ALLILCINPP
501 LLRS
11AA_SEQUENCE 1.0
ID AAG26394 standard; Protein; 329 AA.
XX
AC AAG26394;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 30833.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.
OS
XX
PN EPI033405-A2.
PD
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.

PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 19-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 22-NOV-1999; 99US-0167362.
PR 23-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

XX (CERE-) CERES INC.
PA Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
XX Zheng L, Dumas J;
PI WPI; 2000-507395/46.
XX N-PSDB; AAC41149.
DR
XX
PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
PS Claim 19; SEQ ID 30833; 344pp + CD-ROM; English.
XX
CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 329 AA;
AAG26394 Length: 329 May 30, 2002 09:25 Type: P Check: 3190 ..
1 MVQSQSLSTL TICGVYVSS LLMNRLNSVK ASSLIGDRCV SCQFLRKSPS
51 FRSHWKSILKQ RNLRYEARW PFOGGEGCLD PSSERSSEAN EDILIEFFOL
101 DLATRVQYAM NLEQYDIAQQ LREKLEVEE ESIRLQEGKR GSSAKSEAQD
151 KGISIRLRA DLQNAIDSED YGLAKLRDE ISKLEAESLA VSAKALAFEK
201 AEYAFRLGQK LRHKTFGYRA VVCGMDPICC ESSSMWEAAE VEKLPGRSNO
251 PFYQVLVDVR THPDLVAVV AEDNLFSPK PDKERDHPY ISFLYGGADT
301 AGDFIPVKQL REKYNRPRIHE VPFDSDQED
11AA_SEQUENCE 1.0
ID AAG26395 standard; Protein: 307 AA.
XX
AC AAG26395;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 30834.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.

PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

XX
XX
PA (CERE-) CERES INC.
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
XX WPI; 2000-507395/46.
DR N-PSDB; AAC41149.

PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
XX sequences -

PS Claim 19; SEQ ID 30834; 344pp + CD-ROM; English.
XX
CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 307 AA;

AAG26395 Length: 307 May 30, 2002 09:25 Type: P Check: 4070 ..

1 MNRUNSVKAS SLIGDRVCSC QFLRKSPSFR SHWKS LKORN LLRVEARWPF
51 QGGGEGLDPS SERSESANED ILTFEFDLD ATRVQYAMNL EGYDIAQQLR
101 EKLTEVEEES IRLQEGKRG SAKSEAQDKG ISIIRLRADL QNAIDSEDTG
151 IAAKLRLDEIS KLEAESLAVS AKALAFKAE YAFRLGQKLK HKTFGYRAVV
201 CGMDPICCES SSWMEEAEVE KLPRGSNOPE YQVLVDVPTH PDLVAYVAE
251 DNLFSPEKPD KERFDHPYIS FLYYGADTAG DFIPVKQLRE KYNRPREHVP
301 FDSQDEED

IIAA_SEQUENCE 1.0
ID AAG29842 standard; Protein; 623 AA.

AC AAG29842;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 35574.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164250.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 24-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.

PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

(CERE-) CERES INC.
Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
Zheng L, Dumas J;
WPI; 2000-507395/46.
N-PSDB; AAC42442.

New sequence determined DNA fragments (SDFs) from different plant
species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
protein coding sequences, untranslated regions, or as 3' termination
sequences -

Claim 19; SEQ ID 35574; 344bp + CD-ROM; English.

The present sequence is a putative protein fragment from
Arabidopsis thaliana. Its coding sequence was isolated by carrying out
RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
library which could then be sequenced, allowing the putative protein
sequence(s) to be obtained. This sequence may be useful for protein
identification and for aiding in the elucidation of signal transduction
and metabolic pathways. Its coding sequence has a use in the control of
gene expression as a promoter, coding sequence, 3'UTR or termination
sequence, for controlling the behaviour of a gene within the chromosome,
as a tool for use in genetic mapping, including a use in hybridisation
assays, for recognition or isolation of similar DNA fragments, or for
the identification of a particular organism.


```
XX      Sequence      623 AA;
SQ
AAG29842 Length: 623 May 30, 2002 09:25 Type: P Check: 7775 ..

      1  DEKRSVNGDL RQPHESNMFR CATERDNETN EKSRSVSHFL LCSLXFESIL
51  YCREEDDVE YKERTQCKEK RRMKSTLDSI DALLFSLRA FTSPPAVFVQ
101 IQGCTICLLL ALGMLLAEV RNREVKRIKN SIKAGNSLAF LYQDINLEH
151 SROVKLPRVS VVMPKGFGE HNLHNWRSQI TSLYGGPLEF LFVESTEDP
201 AYHAVSRLLS MYQDHVEAKV VVAGLSTPCS QKIHNLIGV EKMHKDTKYV
251 LFLDDDVRLH PGTIGALTTE MEKNPEVPCS MGFATGRTF FLWGCGMMH
301 ADDEFQDRYG VVSGLRDGY SDDMTLASLA GAHKRLITSP PYAVEPHPLA
351 SDLSFGRYWN YLRKQTEVLE SYISKVNMIM NKALFAVHCY LSWGFEVAPYV
401 MAIHITSAL RIYIKGYHQL EDTTSASGGL SVFLMLHIDR RMRFYVGMML
451 VITFLICTFI ELISMWNLTR REVOLCNMLS PEAPRLSLAT YNWGLPFIIG
501 GHISPMYRGF NKPRTGVEFN VEDYPIPNLY PHTIFQNIKS ALENKHFNVE
551 TPLLIQTVSS SIWSTSLFD RGLIKPSFT SSASWPPEP MYKGLNGKTS
601 LYLFRARIQN IHISNIAMSA RWC

11AA_SEQUENCE 1.0
ID  AAG29843 standard; Protein: 550 AA.
XX
AC  AAG29843;
XX
DT  17-OCT-2000 (first entry)
XX
DE  Arabidopsis thaliana protein fragment SEQ ID NO: 35575.
XX
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
XX  termination sequence.
OS  Arabidopsis thaliana.
XX
PN  EP1033405-A2.
XX
PD  06-SEP-2000.
XX
PF  25-FEB-2000; 2000EP-0301439.
XX
PR  25-FEB-1999; 99US-0121825.
PR  05-MAR-1999; 99US-0123180.
PR  09-MAR-1999; 99US-0123548.
PR  23-MAR-1999; 99US-0125788.
PR  25-MAR-1999; 99US-0126264.
PR  29-MAR-1999; 99US-0126785.
PR  01-APR-1999; 99US-0127462.
PR  06-APR-1999; 99US-0128234.
PR  08-APR-1999; 99US-0128714.
PR  16-APR-1999; 99US-0129845.
PR  19-APR-1999; 99US-0130077.
PR  21-APR-1999; 99US-0130449.
PR  23-APR-1999; 99US-0130510.
PR  23-APR-1999; 99US-0130891.
PR  28-APR-1999; 99US-0131449.
PR  30-APR-1999; 99US-0132048.
PR  04-MAY-1999; 99US-0132407.
PR  05-MAY-1999; 99US-0132484.
PR  06-MAY-1999; 99US-0132485.
PR  06-MAY-1999; 99US-0132486.
PR  06-MAY-1999; 99US-0132487.
```

```
PR  07-MAY-1999; 99US-0132863.
PR  11-MAY-1999; 99US-0134256.
PR  14-MAY-1999; 99US-0134218.
PR  14-MAY-1999; 99US-0134219.
PR  14-MAY-1999; 99US-0134221.
PR  14-MAY-1999; 99US-0134370.
PR  18-MAY-1999; 99US-0134768.
PR  19-MAY-1999; 99US-0134941.
PR  20-MAY-1999; 99US-0135124.
PR  21-MAY-1999; 99US-0135353.
PR  24-MAY-1999; 99US-0135629.
PR  25-MAY-1999; 99US-0136021.
PR  27-MAY-1999; 99US-0136392.
PR  28-MAY-1999; 99US-0136782.
PR  01-JUN-1999; 99US-0137222.
PR  03-JUN-1999; 99US-0137528.
PR  04-JUN-1999; 99US-0137502.
PR  07-JUN-1999; 99US-0137724.
PR  08-JUN-1999; 99US-0138094.
PR  10-JUN-1999; 99US-0138540.
PR  10-JUN-1999; 99US-0138847.
PR  14-JUN-1999; 99US-0139119.
PR  16-JUN-1999; 99US-0139452.
PR  16-JUN-1999; 99US-0139453.
PR  17-JUN-1999; 99US-0139492.
PR  18-JUN-1999; 99US-0139454.
PR  18-JUN-1999; 99US-0139455.
PR  18-JUN-1999; 99US-0139456.
PR  18-JUN-1999; 99US-0139457.
PR  18-JUN-1999; 99US-0139458.
PR  18-JUN-1999; 99US-0139459.
PR  18-JUN-1999; 99US-0139460.
PR  18-JUN-1999; 99US-0139461.
PR  18-JUN-1999; 99US-0139462.
PR  18-JUN-1999; 99US-0139463.
PR  18-JUN-1999; 99US-0139750.
PR  18-JUN-1999; 99US-0139763.
PR  21-JUN-1999; 99US-0139817.
PR  22-JUN-1999; 99US-0139899.
PR  23-JUN-1999; 99US-0140353.
PR  23-JUN-1999; 99US-0140354.
PR  24-JUN-1999; 99US-0140695.
PR  28-JUN-1999; 99US-0140823.
PR  29-JUN-1999; 99US-0140991.
PR  30-JUN-1999; 99US-0141287.
PR  01-JUL-1999; 99US-0141842.
PR  01-JUL-1999; 99US-0142154.
PR  02-JUL-1999; 99US-0142055.
PR  06-JUL-1999; 99US-0142390.
PR  08-JUL-1999; 99US-0142803.
PR  09-JUL-1999; 99US-0142920.
PR  12-JUL-1999; 99US-0142977.
PR  13-JUL-1999; 99US-0143542.
PR  14-JUL-1999; 99US-0143624.
PR  15-JUL-1999; 99US-0144005.
PR  16-JUL-1999; 99US-0144085.
PR  16-JUL-1999; 99US-0144086.
PR  19-JUL-1999; 99US-0144325.
PR  19-JUL-1999; 99US-0144331.
PR  19-JUL-1999; 99US-0144332.
PR  19-JUL-1999; 99US-0144333.
PR  19-JUL-1999; 99US-0144334.
PR  19-JUL-1999; 99US-0144335.
PR  19-JUL-1999; 99US-0144332.
PR  20-JUL-1999; 99US-0144632.
PR  20-JUL-1999; 99US-0144632.
PR  20-JUL-1999; 99US-0144884.
PR  21-JUL-1999; 99US-0144814.
PR  21-JUL-1999; 99US-0145086.
PR  21-JUL-1999; 99US-0145088.
PR  22-JUL-1999; 99US-0145085.
PR  22-JUL-1999; 99US-0145087.
PR  22-JUL-1999; 99US-0145089.
PR  22-JUL-1999; 99US-0145192.
```

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0147038.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.

PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179395.
PR 01-FEB-2000; 2000US-0179398.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX
XX
PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
XX WPI; 2000-507395/46.
DR N-PSDB; AAC42442.
XX
XX
PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
XX
PS Claim 19; SEQ ID 35575; 344pp + CD-ROM; English.
XX
XX
CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 550 AA;

AAG29843 Length: 550 May 30, 2002 09:25 Type: P Check: 6852 ..
1 MSTLDSIDAI LFSISRAFTS PPAVFVQIQG CTICLLALG WLLAEYVRNR
51 EVKRIKNSIK AGNSIAFLYQ DINELEHSRQ VKLPRVSVVM PLKGFGEHNL
101 HNWRSQITSL YGCPLEFLFV VESTEDPAYH AVSRLLSMYQ DHEAKVVVA
151 GLSTTCQSKI HNLIGVEKM HKDTRYVLF DDDVRLHPGT IGALETTEK
201 NPEVPCSMGF ATGGRTEFLW GCGMMMHADD FRQDRYGVVS GLRDGYSDD
251 MTLASLAGAH KRLITSPPA VFPHPLASDL SFGRYWNYLR KQTFVLESYI
301 SKVNMIMNKA LFAVHCYLSW GFVAPYVMAI IHITSALRIY IKGYHLEDT
351 TSASGGLSVF LMLHIDRRMR FVVGMLLVIT LAICTEIELL SMWNLTREV
401 QLCNMLSPEA PRLSLATYMW GLPFIIGHI SPMYRGFNKP RTGVFWNVED
451 YPIPNLYPHT IFQNIKSALE NKHNVEETPL LIQTWSSSIW LSTSLFDRGI
501 LIKPSFTSSA ASWPETMYK GLNGKTSLYL FARAIONIH I SNIAWSARWC
11AA_SEQUENCE 1.0
ID AAG29844 standard; Protein; 461 AA.
XX
AC AAG29844;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35576.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 03-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.

PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181214.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
DR WPI: 2000-507395/46.
DR N-PSDB; AAC42442.
XX
PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
XX
PS Claim 19; SEQ ID 35576; 344pp + CD-ROM; English.
XX
CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 461 AA;

AAG39844 Length: 461 May 30, 2002 09:25 Type: P Check: 4529 ..

1 MPLKGFGEHN LHMRSQITS LYGGPLEFLF VVESTEDPAY HAVSRLLSMY
51 QDHVEAKVVV AGLSTTCSQK IHNQLIGVEK MHRDTRYVLV IDDDVRLHPG

101 TIGALTTEME KNPEVPCSMG FATGRTFFL WGCMMMHAD DFRDRYGVV
151 SGLRDGYSD DMTLASLAGA HKRLITSPV AVFPHPLASD LSEGRYNYL
201 RKQTFVLESY ISKVMWIMNK ALFAVHCYLS WGFVAPYVMA IHHITSALRI
251 YIKGYHOLED TTSASGSLV FLMHIDRRM RFYVGMMLVI TLAICTFIEL
301 LSMWNLTRRE VOLCNMLSPE APRLSLATYN WGLPFIIGH ISPMYRGFNK
351 PRTGVFWNVE DYPINLYPH TIFONIKSAL ENKHFNVEIP LLLQTVSSSI.
401 WLSTSLFDRG ILIKPSFTSS AASWPPEIMY KGLNGKTSLY LFAAIONTH
451 ISNIAMSARW C
11AA_SEQUENCE 1.0
ID AAG36897 standard; Protein; 333 AA.
XX
AC AAG36897;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45282.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 03-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 08-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 03-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.

PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 28-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX
XX
PA (CERE-) CERES INC.
XX
XX Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
XX WPI; 2000-507395/46.
DR N-PSDB; AAC45095.
XX

PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -

PS Claim 19; SEQ ID 45282; 344pp + CD-ROM; English.

XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX

SQ Sequence 333 AA;

AAG36897 Length: 333 May 30, 2002 09:25 Type: P Check: 2209 ..

1 MYSRNSDGYF LDPTGMTVPG LGPSFTAAYS SSSSPPTSGT AVAVADVTAM
51 VSSSEEDLSK KIRKPYTITK SRESWTEPEH DKFLALQLF DRDWKRIEAF
101 IGSKTVIQIR SHAKYFLKV OKSGTGEHLR PPRPKRAAH PYPQAKHKNV
151 QLOVPGSFRS TSEPNDPSFM FRPESSMLM TSPTTAAAP WTNNAQTISF
201 TPLPKAYCCS TLMETAGAGA NNNCSSSEN TPRPRSNRDA RDHGNVGHSL
251 RVLPDFAQVY GFIGSVFDPY ASNHLQKCLK MDPIDVETVL LLMRNLSINL

301 SSPDEFDHHW TLDYKKRCIQ GISNLLDQKN MKT
IIA_SEQUENCE 1.0
ID AAG36898 standard; Protein; 318 AA.
XX
AC AAG36898;
XX
DF 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45283.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.

PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.

PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX
PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
PI
XX
DR WPI; 2000-507395/46.
DR N-PSDB; AAC45095.
XX
PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
PS Claim 19; SEQ ID 45283; 344pp + CD-ROM; English.
XX

CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 318 AA;

AAG36898 Length: 318 May 30, 2002 09:25 Type: P Check: 464 ..

1 MTVPGLGPF TAAVSSSSP TTSTAVAVA DVTAMVSSSE EDLSKKIRKP
51 YTIITKSRESW TEPEHDKFLE ALQLFDRDWK KIEAFIGSKT VIQIRSHAOK
101 YFLKVQKSGT GEHLPPRPK RKAAPYPQK AHKNVQLQVP GSFKSTSEPN
151 DPEFMFRPES SSMIMTSPTT AAAAPWTNNA QTISETPLPK AYCCSTLMET
201 AGAGANNMCS SSENTPRPR SNRDARDHGN VGHSLRVLPD FAQVYGFICS
251 VEDPYASNHL QKLKMPID VETVLLMRN LSINLSSPDE EDHWHITLDYK
301 KRCIQGISNL LDQKNMKT

!!AA_SEQUENCE 1.0
ID AAG36899 standard; Protein; 284 AA.
XX
AC AAG36899;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45284.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.

PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 03-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.

PR 17-FEB-2000; 2000US-0183166.
XX
PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
DR WPI; 2000-507395/46.
DR N-PSDB; AAC45095.
XX
PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
PS Claim 19; SEQ ID 45284; 344pp + CD-ROM; English.
XX
CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 284 AA;
AAG36899 Length: 284 May 30, 2002 09:25 Type: P Check: 6630 ..
1 MVSSEEDLS KRIKPYTIT KRSRWTEPE HDKFLALQL FDRDWKITEA
51 FIGSKTVIOI RSHAQYFLK VQKSGTGEHL PPPRPKRKAA HPIYQAKHKN
101 VOLOVPGSFK STSEPNDSF MRPRESSML MTSPTTAAA PWTNNAQTIS
151 FTPLPKAYCC STLMTAGAG ANNNCSSSE NTPRPNSNRD ARDHGNVGH
201 LRVLPDFAQV YGFIGSVFDP YASNHLQKLK KMDPIDVETV LLLMRNLISN
251 LSSPDFEDHW HTLDYKKRCI QGISNLLDQK NMKT
IIAA_SEQUENCE 1.0
ID AAG37272 standard; Protein; 584 AA.
XX
AC AAG37272;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45804.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.

PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 03-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX
PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
DR WPI; 2000-507395/46.
DR N-PSDB; AAC45242.
XX
PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
PS Claim 19; SEQ ID 45804; 344pp + CD-ROM; English.

XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 584 AA;

AAG37272 Length: 584 May 30, 2002 09:25 Type: P Check: 7852 ..

1 YFEIIVLFL LIRIQNSSF KLFINLVFEF WSIETMGRGS VTSIAPGFRF
51 HPTDEELVRY YLKRKICNKP EKFDALSVTD VYKSEPDWLP DKSRLKSRLD
101 EWFYFSMLDK KYRNGSKTNR ATEMGYWKTT GKDRRELNGS KVGGMKKTLY
151 YHKGRAPRGE RTNWVMHEYR LYDQDLDKTG VHQDAFVLCR IFQKSGSGPK
201 NGEQYGAPFV EEEWEEEDDM TFVPDQEDLG SEDHYVYHMD DIDQSENFV
251 VYDAIPPLN FIGHSSNNV ETNYSDSINY IQQTGNVMS GGYEQAPAES
301 YEKDQPIIR DRDGLONEG IGCYVDKHS ETLQSSDNIF GTDTSYDNF
351 PVESNYLIGE AFLDPNSNLL ENDGLYLETN DLSSTQODGF DFEDYLTFFD
401 ETFDPSQLMG NEDVFPOEE LFOEVETKEL EKEETSRSKH VVEEKEKDEA
451 SCSKQVDADA TEFEPDYKYP LKKASHMLG AIPAPLANAS EFPYKDAIR
501 LHAAOSSGSV HVTAGMTTIS DSNMGWSYCK NENLDLILSL GLVQGNTPAE
551 KSGNSSANAM LIPMCFWVLL LSVSEKVSIL VSSR

IIAA_SEQUENCE 1.0
ID AAG37273 standard; Protein; 549 AA.

XX AAG37273;

XX AC AAG37273;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 45805.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.

XX
SQ Sequence 549 AA;

AAG37273 Length: 549 May 30, 2002 09:25 Type: P Check: 1025 ..

1 MGRGSVTSLA PGFRFHPTDE ELVRYLLKRR ICNKPFFDA ISVTDVYKSE
51 PWDLPDKSRL KSRDLEWYRF SMLDKKYRNG SKTNRATEMG YWKTGKDRE
101 ILNGSKVVGCM KKTLYVHKGR APRGERINWV MHEYRLVDQD LDKTGVMQDA
151 FVLCRIFQKS GSGPKNGEYQ GAPVEEEME EEDDMTFVPD QEDLGSEDHV
201 YVHMDIDQK SENFVYDAI PIPLNFIHGE SSNNVETNYS DSINTYQQTG
251 NYMDSGGYFE QPAESYEKQD KPIIRDQGS LQNEGICGV QDKHSETLOS
301 SDNIFGTDTS CYNDPVEESN YLIGEAFLDP NSNLLDNDGL YLETNDISST
351 QQDGFDFEDY LTFPDETFDP SOLMGNEDEV FDQEELEQEV ETKELEKEET
401 SRSKHVVEEK EKDEASCSKQ VDADATEFERP DYKYPPLKKA SHMLGAIPAP
451 LANASEPPTK DAIRLHAAQ SSGSVHVTAG MITISDSNMG WSYGKNENLD
501 LILSLGLVQG NTAPEKSGNS SAWAMLIFMC FWVLLLSVSF KVSILVSSR

11AA_SEQUENCE 1.0
ID AAG37274 standard; Protein; 478 AA.

XX AC AAG37274;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 45806.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 99US-0132048.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.

PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
PA (CERE-) CERES INC.
XX
XX Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
XX WPI: 2000-507395/46.
DR N-PSDB: AAC45242.
XX
XX New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
XX Claim 19; SEQ ID 45806; 344pp + CD-ROM; English.
PS
XX
XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,

CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 478 AA;
AAG37274 Length: 478 May 30, 2002 09:25 Type: P Check: 969 ..
1 MLDKKYRNGS KTNRATEMGY WKTGKDKREI LNSKVVGMK KTLVYHKGRA
51 PRGERTNWVM HEYRLVDQL DKTGVHODAF VLCRIFOKSG SGPKNGEQYG
101 APEVEEEMEE EDDMTFVPDQ EDLGESEHVV VMDDIDQKS ENFVYDAIP
151 IPLNFIHGES SNNVETNYSI SINYIQOTGN YMDSGGYFEQ PAESYEKDK
201 PIIRDRDGLS QNEGICGCVQ DKHSETLQSS DNIFGTDTSC YNDFPVESNY
251 LIGEAFLDPN SNLENDGLY LETNDLSSTQ QDGFDEFDYL TFFDETDPDS
301 QLMGNEDVFF DQELFQVEE TKELEKEETS RSKHVEEKE KDEASCSKQV
351 DADATEFEPD YKPYLLKKAS HMLGAIPAPL ANASEPTKD AAIRLHAQGS
401 SGSVHTAGM ITISDSNMGW SYCKNENLDL ILSLGLVQGN TAPEKSGNSS
451 AWAMLIEMCF WVLLLSVSFK VSILVSSR
11AA_SEQUENCE 1.0
ID AAG38413 standard; Protein; 261 AA.
AC AAG38413;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47387.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163249.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.

PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 04-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181214.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181476.
PR 15-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

(CERE-) CERES INC.

PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;

DR WPI: 2000-507395/46.
DR N-PSDB; AAC45684.

PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -

PS Claim 19; SEQ ID 47387; 344pp + CD-ROM; English.

CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.

XX Sequence 261 AA;

AAAG38413 Length: 261 May 30, 2002 09:25 Type: P Check: 3958 ..

1 MWQSQSLSLT TICGSVKVSS LLRNRLNSVK ASSLIGDRCV SCQFLRKSPS
51 FRSHWKSILKQ RNLLRVEARW PFOGGEGQGL DPSSERSESA NEDILIFFFQ
101 LDLATRVQYA MNLEQYDIAQ QLREKLTVEE ESIRLQEGK RGSSEAKSEAQ
151 DKGISIRLR ADLQNAIDSE DYGLAKLRD EISKLEAEST AVSAKALAFE
201 KAEYAFRLGQ KLRHKTFGYR AVVCGMDPIC SSSSSWMEAA EVEKLPRGSN
251 QPFYQLKTI C
11AA_SEQUENCE 1.0
ID AAG38839 standard; Protein; 986 AA.
XX
AC AAG38839;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47973.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.

PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.

PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX
XX
PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
XX WPI; 2000-507395/46.
DR N-PSDB; AAC45844.
XX
XX

PT New sequence determined DNA fragments (SDEs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
XX
XX

PS Claim 19; SEQ ID 47973; 344pp + CD-ROM; English.
XX
XX

CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
XX

SQ Sequence 986 AA;

AAG38839 Length: 986 May 30, 2002 09:25 Type: P Check: 7690 ..

1 MGRESVAVVT APPSATAPGT ASVATSLAPG FRHPTDEEL VSYLKRKVL
51 GQPVREDAIG EVDIYKHEPW DLAVFSRLKT RDQEWYFYSA LDKKYNGAR
101 MNRATNRGYW KATGKDREIR RDILLGMRK TLVHSGRAP DGLRTNWVMH
151 EYRLVEYTE KNGNLVQDAY VLCRVFHKN IGPPSGNRYA PFMEEWADD
201 EGALLPGIDV KLRLEPPVA NGNDQMDQEI QSASKSLINI NEPPREAPL
251 DIESDQNNHH ENDLKPEEHN NNNNYDENEE TLKREQMEEE ERPPRPVCL
301 NKEAPLPLQ YKRRQSESN NNSRNTQDH CSSTTTVDN TTLLISSAA
351 ATNTAISALL EFSLMGISDK KEKPOPLRP HKEPLPPQTP LASPEEKRS

401 VCYFELIVLL FLIRIRONS SEKLEINLVF EFWSEIYMR GSVTSLAPG
451 RFHPTDEELV RYLYKRKICN KPEKDAISV TDVYKSEPWD LPDKSRLKSR
501 DLEWYFESML DKRYNGSKT NRATEMGYWK TTGKDREILN GSKVVGMRKT
551 LVYHKGRAPR GERTNWVME YRLVDQDLK TGVHODAFVL CRIFOKSGSG
601 PKNGEYQAP FVEEWEED DMTFVPDQED LGSEDHVYVH MDDIDOKSEN
651 FVYVDAIPIP LNFTHGESSN NVETNYSDSI NYIQGTGNYM DSGGYFEQPA
701 ESEKDKPI IRPBDGSLON EGIGCGVQDK HSETLQSSDN IFGTDTCYN
751 DFPVESNYLI GEAFLDPSN LLENDGLYLE TNDLSSTQOD GFDFEDYITF
801 FDETFDPSQL MGNEDVFFDQ EELFQEVETK ELEKEETSRG KHVEEKED
851 EASCSKQVDA DATEFEPDYK YPLIKKASHM LGAIAPAPLAN ASEPTKDA
901 IRLHAAQSSG SVHTAGMIT ISDSNMGWSY GRNENLDLIL SLGLVQNTA
951 PEKSGNSSAW AMLIFMCFWV LLSVSFKVS ILVSSR

!!AA_SEQUENCE 1.0
ID AAG38840 standard; Protein; 886 AA.
XX
XX

AC AAG38840;
XX

DT 18-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 47974.
XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX

OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX

PD 06-SEP-2000.
XX

PF 25-FEB-2000; 2000EP-0301439.
XX

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.

PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX
XX
PA (CERE-) CERES INC.
XX
XX Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
PI
XX
XX WPI; 2000-507395/46.
DR N-PSDB; AAC45844.
XX
XX
XX
PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
XX
PS Claim 19; SEQ ID 47974; 344bp + CD-ROM; English.
XX
XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
XX
SQ Sequence 886 AA;
AAG38840 Length: 886 May 30, 2002 09:25 Type: P Check: 7810 ..

1 MNRATNRCYW KATCKDREIR RDILLGMMK TLVFSGRAP DGLRTNVMH
51 EYRLVEYTE KNGNLVQDAY VLCRVFFHKNN IGPPSGNRYA PFMEREWADD

101 EGALIPGIDV KLRLEPPVA NGNDQMDQEI QSASKSLINI NEPPRETAPL
151 DIESDQQNH ENDLKPEEHN NNNNYDENEE TLKREQMEEE ERPPRPVCVL
201 NKRAPLPLQ YKRRQSESN NNSSRNTQDH CSSTTTVDN TTTLISSAA
251 AINTAISALL EPSIMGISDK KEKPOQLRP HKEPLPPQTP IASPEEKRS
301 VCYEETIIVL FLIRIRQNS SPKLEINLVE EFWSIEYMR GSVTSLAPGF
351 RPHPTDEELV RYILKRKICN KPEKDAISV TDYKSEPWD LPDKSRLKSR
401 DLEWYFFSML DKYRNGSKT NRATEMGYWK TTGKDREILN GSKVGMKKT
451 LVYHKGRAPR GERTNWMHE YRLVDQDLK TGVHODAFVL CRIFOKSGSG
501 PRNGEQYGAP FVEEWEED DMTFVPDQED LGSEDHVYVH MDDIDQKSEN
551 FVYVDAIPIP LNFHGESEN NVETNYSDSI NYIQQTGNM DSGGYFEQPA
601 ESYEKDQKPI IRDRDGLQN EGIGCGVQDK HSETLQSSDN IFGTDTSQYN
651 DFPVESNYLI GEAFLDPNNS LLENDGLYLE TNDLSSTQOD GFDEFDYLTF
701 FDETFDPSQL MGNEDVFFDQ EELEQEVETK ELEKETSR S KHVEEKED
751 EASCSKQVDA DATEFEPDYK YPLLKASHM LGAIAPPLAN ASEPTKDAA
801 IRLHAAQSSG SVHYTAGMIT ISDSNMGSY GRNENLDLIL SLGLVQNTA
851 PERSGNSSAW AMLIFMCFWV LLSVSFKVS ILVSSR
!!AA_SEQUENCE 1.0
ID AAG38841 standard; Protein; 859 AA.
XX AC AAG38841;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 47975.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PE 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160960.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 10-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165919.
PR 17-NOV-1999; 99US-0165918.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 30-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.

PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX
PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
DR WPI; 2000-507395/46.
DR N-PSDB; AAC45844.
XX
XX
PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
PS Claim 19; SEQ ID 47975; 344bp + CD-ROM; English.
XX
XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX

SQ Sequence 859 AA;
AAG38841 Length: 859 May 30, 2002 09:25 Type: P Check: 4753 ...
1 MKKTLVFHSG RAPDGLRTNW VMHEYRLVEY ETEKNGNLVQ DAYVLCRVFH
51 KNNIGPPSGN RYAPFMEEW ADDEGALIPG IDVKLRLEPP PVANGNDQMD
101 QEIQSASKSL ININEPPRET APDIIESDQQ NHHENDLKPE EHHNNNNYDE
151 NEETLKREQM EEEERPRPV CVLNKEAPLP LIQYKRRQES ESNNNSSRNT
201 QDHCSSTTTT VDNNTTLLISS SAAATNTAIS ALLEFSLMGI SDKKEKPOQP
251 LRPHKEPLPP QTPPLASPEEK RKSVCYFEII VLLFLIRIR QNSSEKLFIN
301 LVPEFWSTIEY MGRGSVTSLA PGFRFHPTE ELVRYYLKRK ICNKPFKFDA
351 ISVTDVYKSE PWDLPDKSRL KSRDLEWYFF SMLDKKYRNG SKTNRATEMG
401 YWKTGKDRE ILNGSKVGM KKTLYVHKGR APRGERTNVV MHEYRLVDQD
451 LDKTGVHQA FVLCRIFQKS GSGPKNGEQY GAPFVEEWE EEDDMTFVPD
501 QEDLGSEDHV YVHMDIDQK SENEVYDAI PIPLNFHGE SSNNVETNYS
551 DSINYIQQTG NYMDSGGEFE QPAESYEKDQ KPIIRDQGS LQNEGIGCV
601 QDKHSETLQS SDNIFGTDTS CYNDFPVESN YLIGEAFLDP NSNLLENDGL
651 YLETNDLST QODGFDFEDY LTFEDETDP SQLMGNEVF FDQELFQEV
701 EFKELEKEET SRSKHVVEEK EKDEASCSKQ VDADATEFEF DYKYPPLKKA
751 SHMLGAIPAP LANASEFPTK DAAIRLHAAQ SSGSVHTAG MITISDSNMG
801 WSYGKNENLD LILSLGLVOG NTAPEKSGNS SAMAMLIIFMC FWVLLLSVSF
851 KVSILVSSR
!!AA_SEQUENCE 1.0
ID AAY95439 standard; Protein; 1503 AA.
AC AAY95439;
XX
XX 10-OCT-2000 (first entry)
DT
XX
DE Human calcium channel polypeptide.
XX
XX Human; SOC-2/CRAC-1; calcium channel; store operated channel;
KW calcium release activated channel; therapy; diagnosis;
KW lymphocyte proliferative disorder.
XX
OS Homo sapiens.
XX
XX WO200040614-A2.
PN
XX
PD 13-JUL-2000.
XX
XX 20-DEC-1999; 99WO-US29996.
PF
XX
XX 30-DEC-1998; 98US-0114220.
PR 29-JAN-1999; 99US-0120018.
PR 22-JUN-1999; 99US-0140415.
XX
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI Scharenberg AM;
XX
XX WPI; 2000-465957/40.
DR N-PSDB; AAA49926.
XX
PT New SOC/CRAC calcium channel polynucleotides and polypeptides used to

PT diagnose and treat proliferative disorders associated with the channel,
PT and to screen for novel modulators of the channel -

XX Example; Page 69-72; 108pp; English.

XX
CC The present sequence is that of a human calcium channel polypeptide
CC as deduced from a cDNA clone (see AAA49926) identified in an EST
CC database search for sequences showing homology to Caenorhabditis
CC elegans calcium signal sequences. Clones isolated from such
CC screenings were used to identify human clones (see AAA49922-24)
CC encoding members (see AAY95435-37) of a new family of SOC (store
CC operated channel) or CRAC (calcium release activated channel)
CC calcium channel polypeptides. SOC/CRAC polypeptides modulate Ca2+
CC flux into and out of a cell, and may be activated upon depletion
CC of Ca2+ from intracellular calcium stores, allowing Ca2+ influx
CC into a cell. SOC/CRAC polypeptides and polynucleotides can
CC be used to treat patients that require modulation of calcium influx
CC into their SOC/CRAC-expressing cells. Polypeptides can be used to
CC deliver therapeutic and/or imaging agents to such cells to modulate
CC proliferation and growth. SOC/CRAC polypeptides also represent
CC targets for designing and/or identifying inhibitors that block
CC lymphocyte proliferation and binding agents that selectively bind
CC to SOC/CRAC polypeptides to which drugs or toxins can be conjugated
CC for delivery to SOC/CRAC expressing cells.

XX
SQ Sequence 1503 AA:

AAY95439 Length: 1503 May 30, 2002 09:25 Type: P Check: 5055 ..

1 MEPSALRKAG SEQEGFEGF PRRVTDLGMV SNLRRSNSL FKSRLQCPF
51 GNNDKQESLS SWIPENIKK ECVYFVSSK LSDAGKVCQ CGYTHEQHLE
101 EATKPHTEFG TQWDPKKHVQ EMPDFAFDI VFTGLSQVK KYVRVSQDTP
151 SSVIYHMTQ HWGLDVPNL ISVTGAKNF NMRPLKSIF RRLGVKVAQT
201 TGAWITGGS HTGVWKQVGE AVRDESLSS YKEGELITIG VATWGTVHRR
251 EGLIHPTGSF PAEYILDEDG QGNLTCLDSN HSHFILVDDG THGQYGEIIP
301 LRTPLEKFI EGTKEGVA IKIPIVCVL EGGPGTLHTI DNATTNGTPC
351 VVVEGSGRVA DVIAQVANLP VSDITISLIQ QKLSVEFOEM FETFTESRIY
401 EWTKKIQDIV RRROLLTVFR EKGQGDQDQD VALIQALLKA SRSQDHFGHE
451 NWDHQLKLV AMNRVDIARS EIFMDEWQWK PSDLHPTMTA ALISNKPEFV
501 KLFLNGVOL KEFVTWDTLL YLYENLDPSC LFHSKLQKVL VEDPERPACA
551 PAAPRLQMHV VAQVLRLLG DFTQPLYP RP RHNDRLRLLL PVPVHKLNVQ
601 GVSLSLYKR SSGHYTFTMD PIRDLLIMAI VQNRRELAGI IMAQSQDCIA
651 AALACSKILK ELSKEEDTD SSEEMALAE EYEHRAIGVF TECYRKDEER
701 AOKLLTRVSE AMGKTTCLQL ALEAKDMKFV SHGCIQAFLT KVMWQLSVD
751 NGLMRYTLCM LAFPLLTLGL ISFREKRLQD VGTPAARARA FFTAPVVVFH
801 LNIISYFAFL CLFAVYLVMD FQVPSWCEC AIYLMFLSV CEEMRQLFYD
851 PDECGLMKKA ALYPSDFWNK LDVGAILLFV AGLCRLIPA TLYPGRVILS
901 LDFILECLRL MHIFTISKTL GPKLIIVKRM MKDYEFFELF LAVWVVSFGV
951 AKQAILIHNE RRVDMLEFGA VYHSYLTIFG QIPGYIDGVN FNDHCSPNG
1001 TDPYKPKCPE SDATQQRPAF PEMLTVLLLC LYLLFTNILL LNLIIAMFNV
1051 TFGQVQEHND QIMKFORHDL IEEYHGRPA PPPFILLSHL QLFIKRVVLK

1101 TPAKRHKQLK NKLKNEEAA LLSWEIYLKE NYLONROFOQ KORPEOKIED
1151 ISNKVDAMVD LLDLDPLKRS GSMQORLASL EEOVAQTARA LHMIVRTLRA
1201 SGFSSEADV PTLASQKAAEE PDAEPGGRKK TEERPGDSYHV NARHLLYPNC
1251 PVTRFPVNE KVPWETEFLI YDPPFYTAER KDAAMDPMG DTLEPLSTIQ
1301 YNVVDGLRDR RSFHGPYTVQ AGLPLNPMGR TGLRGGSLS CFCPNHTLYP
1351 MVRMRNED GAICRKSICK MLEVLYVKLP LSEHWALPGC SREPGEMLP
1401 KLKRIHQEH WPSFENLLKC GMEYKGYMD DPRNTDNAMI ETYAVSVHFO
1451 DQNDVELNRL NSNLHACDSG ASIRWQVDR RIPLYANHKT LLOKAAAEFG
1501 AHY

11AA_SEQUENCE 1.0
ID AAY66728 standard; protein; 280 AA.

XX AC AAY66728;

XX DT 05-APR-2000 (first entry)

XX DE Membrane-bound protein PRO1131.

XX KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
XX KM pharmaceutical; receptor immunoadhesin; gene mapping.

XX OS Homo sapiens.

XX PN WO9963088-A2.

XX PD 09-DEC-1999.

XX PF 02-JUN-1999; 99WO-US12252.

XX PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 04-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 10-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.

ID AAY76218 standard; Protein; 161 AA.
XX
AC AAY76218;
XX
DT 23-MAR-2000 (first entry)
XX
DE Human secreted protein encoded by gene 95.
XX
KW Human; secreted protein; cancer; tumour; developmental abnormality;
KW foetal deficiency; blood disorder; immune system disorder; inflammation;
KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
KW therapy; chromosome 1.
XX
OS Homo sapiens.
XX
PN WO9958660-A1.
XX
PD 18-NOV-1999.
XX
PF 06-MAY-1999; 99WO-US09847.
XX
PR 12-MAY-1998; 98US-0085093.
PR 12-MAY-1998; 98US-0085094.
PR 12-MAY-1998; 98US-0085105.
PR 12-MAY-1998; 98US-0085180.
PR 18-MAY-1998; 98US-0085906.
PR 18-MAY-1998; 98US-0085920.
PR 18-MAY-1998; 98US-0085921.
PR 18-MAY-1998; 98US-0085922.
PR 18-MAY-1998; 98US-0085923.
PR 18-MAY-1998; 98US-0085924.
PR 18-MAY-1998; 98US-0085928.
PR 18-MAY-1998; 98US-0085925.
PR 18-MAY-1998; 98US-0085927.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
PI Lafleur DW, Endress GA, Ebner R;
XX
DR WPI; 2000-062296/05.
DR N-PSDB; AAZ65344.
XX
PT New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -
XX
PS Claim 11; Page 412; 475pp; English.
XX
CC AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
CC AAY76124 to AAY76223 are the secreted proteins encoded by the 97 human
CC genes. The gene encoding this protein was found to be on chromosome 1.
CC The genes and their corresponding secreted polypeptides are
CC useful for preventing, treating or ameliorating medical conditions,
CC e.g. by protein or gene therapy. Also pathological conditions can be
CC diagnosed by determining the presence of mutations in the new polypeptides in a sample
CC or by determining the presence of mutations in the new genes. Specific
CC uses are described for each of the 97 genes, based on which tissues they
CC are most highly expressed in, and include developing products for the
CC diagnosis or treatment of cancer, tumours, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
CC disorders, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The sequences shown in AAY76224 to AAY76424 represent fragments of the
CC secreted proteins.
XX

SO Sequence 161 AA;
AAY76218 Length: 161 May 30, 2002 09:25 Type: P Check: 8440 ..
1 MAFTFAAFCY MSLVLCAAL IFFAIWHIIA FDELRTDFKS PIDQCNPVHA
51 RERLENIERI CFLRLKVLVP EYSIHSLEFCI MFLCAQEWLT LGINVPPLFF
101 HFWRYFHCPA DSSEIAYDPP VVMNADTLSY CQKEAXCKLA FYLLSFFYYL
151 YCMITLVSS X
11AA_SEQUENCE 1.0
ID AAY53783 standard; Protein; 541 AA.
XX
AC AAY53783;
XX
DT 22-FEB-2000 (first entry)
XX
DE A maize phosphate transporter designated Zmpt1.
XX
KW phosphate transporter; zmpt1; phosphorous uptake; phosphate poor soil;
KW free phosphorus; animal consumption; seed; phosphate level; germination.
XX
OS Zea mays.
XX
PN WO9958657-A2.
XX
PD 18-NOV-1999.
XX
PF 11-MAY-1999; 99WO-US10374.
XX
PR 13-MAY-1998; 98US-0085292.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI McEliver JA, Zhao S, Bowen BA, Wang X;
XX
DR WPI; 2000-053096/04.
DR N-PSDB; AAZ36416.
XX
PT Zea mays phosphate transporter genes and proteins, transgenic plants
PT and seeds -
XX
PS Claim 9; Page 86-88; 108pp; English.
XX
CC The present sequence represents a Zea mays phosphate transporter,
CC designated zmpt1. The polynucleotide sequence was identified and isolated
CC from a Zea mays nucleic acid library. The phosphate transporter
CC polynucleotides and polypeptides are useful for altering phosphate
CC transporter expression in transgenic plants. They are useful for
CC providing plants capable of increasing phosphorous uptake in phosphate
CC poor soils, or to increase the level of free phosphorus in plants used
CC for animal consumption, or to decrease the level of free phosphate in
CC seeds where phosphate levels would otherwise be deleterious to
CC germination. The invention also provides means for identifying compounds
CC that bind to and/or increase or decrease the enzymatic activity of the
CC zmpt polypeptides.
XX
SO Sequence 541 AA;
AAY53783 Length: 541 May 30, 2002 09:25 Type: P Check: 6982 ..
1 MARGGGLQV LSAIDAAKTQ WYHETAIIVA GMSGFTDAYD LFCISLVTKL
51 LGRITYTDTs KDNPGSLPPN VAAAVNGVAF CGTLAQGLFF GMLGDKLGRK
101 SVYGMULMIM VICSVASGLS FGHPTGVMA TLCEFFRWLG FGIGGDYPLS
151 ATIMSEYANK KTRGAFIAAV FAMQFGILA GGIVTLVISA AFRAGYPAPA
201 YRQDHFNSTV PQADYVWRII LILGAAPAML TYTWKMKPE TARYTALVAK

251 NAKQAAADMS RVLQTEIVDE QEKIDEMVTA ESNTFGLFSR EFARRHGLHL
301 VGTSTTWELL DIAFYSQNLF QKDIETSIW IPKANTMSAL EEVERISRQ
351 TLIALCGTVP GYWEFTVALID VVGRAIQLL GFEFMTVEML GLAIPYHHWT
401 TPGNHIGFV MYAFTFFFN EGPNSTTFIV PAEIFPARLR STCHGISAAS
451 GKAGAIIGAF GFLYAAQND KSKADAGYPA GIGVRNSLFV LAASNILGFI
501 LTFVLPEESKG KSLSEMSGEA DDAEDDAVGT RAVRPSGTQM V

11AA_SEQUENCE 1.0
ID AAY53784 standard; Protein; 542 AA.

AC AAY53784;
XX
DT 22-FEB-2000 (first entry)

DE A maize phosphate transporter designated ZmptII.

KW Phosphate transporter; ZmptII; phosphorous uptake; phosphate poor soil;
KW free phosphorus; animal consumption; seed; phosphate level; germination.

XX Zea mays.

PN W09958657-A2.

PD 18-NOV-1999.

PF 11-MAY-1999; 99WO-US10374.

PR 13-MAY-1998; 98US-0085292.

PA (PION-) PIONEER HI-BRED INT INC.

PI McElver JA, Zhao S, Bowen BA, Wang X;

DR WPI; 2000-053096/04.

DR N-PSDB; AAZ36417.

PT Zea mays phosphate transporter genes and proteins, transgenic plants

PS Claim 9; Page 90-91; 108pp; English.

CC The present sequence represents a Zea mays phosphate transporter,
CC designated ZmptII. The polynucleotide sequence was identified and
CC isolated from a Zea mays nucleic acid library. The phosphate transporter
CC polynucleotides and polypeptides are useful for altering phosphate
CC transporter expression in transgenic plants. They are useful for
CC providing plants capable of increasing phosphorous uptake in phosphate
CC poor soils, or to increase the level of free phosphorus in plants used
CC for animal consumption, or to decrease the level of free phosphate in
CC seeds where phosphate levels would otherwise be deleterious to
CC germination. The invention also provides means for identifying compounds
CC that bind to and/or increase or decrease the enzymatic activity of the
CC Zmpt polypeptides.

XX Sequence 542 AA;

AAY53784 Length: 542 May 30, 2002 09:25 Type: P Check: 8403 ..

251 NAKQAAADMS RVLHTEILDE QEKIDGMVAA EGANSFGLFS REFARRHGLH
301 LVGTATTWFL LDIAFYSQNL FQKDIETSIW WIPKANTMSA LEEVYRISRA
351 QTLIALCGTV PGYWEFTVALI DVVGRAIQL LGFEFMTVEM LGLAIPYHHW
401 TTPGNHIGFV VMYAFTFFFA NEGPNSTTFI VPAEIFPARL RSTCHGISAA
451 SKAGAIIGA FGFLYAAQND DSKTDAGYP AGIGVRNSLF VLAASNMLGF
501 VLTFVLPEESK GKSLEMSG EADSEEPVG ARAVRPSETQ MV

11AA_SEQUENCE 1.0
ID AAY53785 standard; Protein; 547 AA.

AC AAY53785;
XX
DT 22-FEB-2000 (first entry)

DE A maize phosphate transporter designated ZmptIII.

KW Phosphate transporter; ZmptIII; phosphorous uptake; phosphate poor soil;
KW free phosphorus; animal consumption; seed; phosphate level; germination.

XX Zea mays.

PN W09958657-A2.

PD 18-NOV-1999.

PF 11-MAY-1999; 99WO-US10374.

PR 13-MAY-1998; 98US-0085292.

PA (PION-) PIONEER HI-BRED INT INC.

PI McElver JA, Zhao S, Bowen BA, Wang X;

DR WPI; 2000-053096/04.

DR N-PSDB; AAZ38418.

PT Zea mays phosphate transporter genes and proteins, transgenic plants

PS Claim 9; Page 94-95; 108pp; English.

CC The present sequence represents a Zea mays phosphate transporter,
CC designated ZmptIII. The polynucleotide sequence was identified and
CC isolated from a Zea mays nucleic acid library. The phosphate transporter
CC polynucleotides and polypeptides are useful for altering phosphate
CC transporter expression in transgenic plants. They are useful for
CC providing plants capable of increasing phosphorous uptake in phosphate
CC poor soils, or to increase the level of free phosphorus in plants used
CC for animal consumption, or to decrease the level of free phosphate in
CC seeds where phosphate levels would otherwise be deleterious to
CC germination. The invention also provides means for identifying compounds
CC that bind to and/or increase or decrease the enzymatic activity of the
CC Zmpt polypeptides.

XX Sequence 547 AA;

AAY53785 Length: 547 May 30, 2002 09:25 Type: P Check: 6498 ..

1 MARGDGLQV LSALDAKTQ WYHFTAIIVA GMGFTDAYD LFCISLVTKL
51 LGRITYTDTS KDSPGSLPPN VAAAVNGVAF CCTLGQLTF GWLGDKLGRK
101 SVYGMTLMVM VICSVASGLS FGHPTGVMA TLCFFRFLWG FGIGDYPIS
151 ATIMSEYANK RTRGAFAIAV FAMQGFGLA GGIIVTLVISA AFRAAYPSDA
201 YRDDHTSTV PQADFVWRVI LMLGAAPALL TTYWRMKMPE TARYTALVAK

1 MAAGDLEVLT ALDTAKTQWY HFTAIIVAGM GFTDAYDLF CISLVTKLIG
51 RIYTVESGA TPGTLPPHVS ASVNGVAFVG TLSGQLFFGW LGDKLGRKV
101 YGMTLMVLV CSVASGLSFG HTPASVMATL CFFRFLWFG IGDYPLSAT
151 IMSEYANKKT RGAFTIAVFA MQGFIMAGC LVAIVSAMP KASFPAPAYA
201 VDPASTPPQ ADFVWRIIM LGAMPALTY YWRTKMPETA RYTALVAKNA

251 KQAAADMSKV LQVEISAGAA EEGAAAATAT ATEPAPASAS ASFGLFSGEF
301 LRRHGLHLIG TSLTWFLLDI AFYSONLFQK DIFSAGWIP KAATMALBEE
351 LFSIARAQSL IALCGTVPGY WFTVALIDVL GRFAIQVTGF LMMTVFMLGL
401 AVPEYHWTTP GHHIGFIVMY GLTEFFANFG PNATTFIVPA EIFPARLRST
451 CHGISAASGK LGAIIISFGF LYLAQNRPDA KTDHGYPAGI GVRNSLFLLA
501 GCNLLGLAFT FLVPESKGKS LEEMSGENDE AAAAAATPNY NNRTVPV

11AA_SEQUENCE 1.0
ID AAY53786 standard; Protein; 539 AA.

AC AAY53786;

DT 22-FEB-2000 (first entry)

DE A maize phosphate transporter designated ZmptIV.

KW Phosphate transporter; ZmptIV; phosphorous uptake; phosphate poor soil;
KM free phosphorus; animal consumption; seed; phosphate level; germination.

OS Zea mays.

PN WO9958657-A2.

PD 18-NOV-1999.

PF 11-MAY-1999; 99WO-US10374.

PR 13-MAY-1998; 98US-0085292.

PA (PION-) PIONEER HI-BRED INT INC.

PI McElver JA, Zhao S, Bowen BA, Wang X;

DR WPI; 2000-053096/04.

DR N-PSDB; AA236419.

PT Zea mays phosphate transporter genes and proteins, transgenic plants
and seeds

PS Claim 9; Page 98-99; 108pp; English.

CC The present sequence represents a Zea mays phosphate transporter,
CC designated ZmptIV. The polynucleotide sequence was identified and
CC isolated from a Zea mays nucleic acid library. The phosphate transporter
CC polynucleotides and polypeptides are useful for altering phosphate
CC transporter expression in transgenic plants. They are useful for
CC providing plants capable of increasing phosphorous uptake in phosphate
CC poor soils, or to increase the level of free phosphorus in plants used
CC for animal consumption, or to decrease the level of free phosphate in
CC seeds where phosphate levels would otherwise be deleterious to
CC germination. The invention also provides means for identifying compounds
CC that bind to and/or increase or decrease the enzymatic activity of the
CC Zmpt polypeptides.

SQ Sequence 539 AA;

AAY53786 Length: 539 May 30, 2002 09:25 Type: P Check: 3508 ..

1 MARGGDGLQV LSALDAKTQ WYHETAIIVA GMGFTDAYD LFCISLVTKL
51 LGRIYYTDT S KDSPGSLPN VAAAVNGVAF CGTLAGOLFF GMLGDKLGRK
101 SVYGMTLMVM VICSVASGLS FGHPTGVMA TLCEFRFWLG FGIGGDYPLS
151 ATIMSEYANK PTRGAFIAAV FAMQGFILA GCIIVLVISA AFRAAYPSPA
201 YRDDHFTSTV PQADIVRVVI VMLGAAPALL TYWLMKMP E TARYTALVAK

251 NAKQAAADMS KVLHTEIVDE QEKIDAAEGA NSFGLFSREF ARRHGLHLVG
301 TATTWFLLDI AFYSONLFQK DIFTSINWIP KANTMSALBEE VYRISRAQTL
351 IALCGTVPGY WFTVALIDVV GRFAIQLLGF FMTVFMGLL AIPYHWTTP
401 GNHIGFVVMY AFTEFFANFG PNSTTFIVPA EIFPARLRST CHGISAASGK
451 AGAIGAFGE LYAAQNDRS KTDAGYPAGI GVRNSLFLVA ASNMLGFVLT
501 FLVPESKGKS LEEMSGEAD SEEPYVGARA VRPSETQMV

11AA_SEQUENCE 1.0
ID AAY53787 standard; Protein; 509 AA.

AC AAY53787;

DT 22-FEB-2000 (first entry)

DE A maize phosphate transporter designated ZmptV.

KW Phosphate transporter; ZmptV; phosphorous uptake; phosphate poor soil;
KM free phosphorus; animal consumption; seed; phosphate level; germination.

OS Zea mays.

PN WO9958657-A2.

PD 18-NOV-1999.

PF 11-MAY-1999; 99WO-US10374.

PR 13-MAY-1998; 98US-0085292.

PA (PION-) PIONEER HI-BRED INT INC.

PI McElver JA, Zhao S, Bowen BA, Wang X;

DR WPI; 2000-053096/04.

DR N-PSDB; AA236420.

PT Zea mays phosphate transporter genes and proteins, transgenic plants
and seeds

PS Claim 9; Page 101-102; 108pp; English.

CC The present sequence represents a Zea mays phosphate transporter,
CC designated ZmptV. The polynucleotide sequence was identified and
CC isolated from a Zea mays nucleic acid library. The phosphate transporter
CC polynucleotides and polypeptides are useful for altering phosphate
CC transporter expression in transgenic plants. They are useful for
CC providing plants capable of increasing phosphorous uptake in phosphate
CC poor soils, or to increase the level of free phosphorus in plants used
CC for animal consumption, or to decrease the level of free phosphate in
CC seeds where phosphate levels would otherwise be deleterious to
CC germination. The invention also provides means for identifying compounds
CC that bind to and/or increase or decrease the enzymatic activity of the
CC Zmpt polypeptides.

SQ Sequence 509 AA;

AAY53787 Length: 509 May 30, 2002 09:25 Type: P Check: 2070 ..

1 MALGQQLRVL HALDVARTQL YHEMATAIAG MGFFTDAYDF FAISLYMDLI
51 SYLYNEQID RGVKATINGI ALGCAVPGQL VFGWLGDKMG RKRIYGVTL
101 LMVYTSIASG LYFGTNEASN VVAVLCEFRF WLFGIGIGDY PLSATIMSEY
151 ANKRRRGAFI AAVFAMQGF NLAAIVAVV VSASFRLRNP RRNANFWMRI
201 VMLGAVPAI LTYWRMKMP ETARYTALVA KDARKASDM SSVLHVEIIP

251 EDEAVRQDKY GLFSAQFLRY HGTHLATSA CWLAVDITFY SLNLYMKDIF
301 ADVGLIDPPG NNDLFRTMV TTLHTGIAL CGTLPGYFTT VAFVDRIGRV
351 RIQLGFTMM SVLTALLAAT YAYWKROETI QRKMGFAVLY GLTNEFANFG
401 PNTTFIVPA EIFPARMRAT CHGIAGARGK IGAIGVFGF MSNMEHGVV
451 PRKLWALPAS NLVGLVFTEL LPDSKGSLE EMAGTEEQO QOQDAVVAA
501 ADHINLVPI
11AA_SEQUENCE 1.0
ID AAY53788 standard; Protein; 554 AA.
XX
AC AAY53788;
XX
DT 22-FEB-2000 (first entry)
XX
DE A maize phosphate transporter designated ZmptVI.
XX
KW Phosphate transporter; ZmptVI; phosphorous uptake; phosphate poor soil;
XX free phosphorus; animal consumption; seed; phosphate level; germination.
XX
OS Zea mays.
XX
PN WO9958657-A2.
XX
PD 18-NOV-1999.
XX
PF 11-MAY-1999; 99WO-US10374.
XX
PR 13-MAY-1998; 98US-0085292.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI McElver JA, Zhao S, Bowen BA, Wang X;
XX
DR WPI; 2000-053096/04.
XX
DR N-PSDB; AA236421.
XX
PT Zea mays phosphate transporter genes and proteins, transgenic plants
PT and seeds -
XX
PS Claim 9; Page 105-106; 108pp; English.
XX
CC The present sequence represents a Zea mays phosphate transporter,
CC designated ZmptVI. The polynucleotide sequence was identified and
CC isolated from a Zea mays nucleic acid library. The phosphate transporter
CC polynucleotides and polypeptides are useful for altering phosphate
CC transporter expression in transgenic plants. They are useful for
CC providing plants capable of increasing phosphorous uptake in phosphate
CC poor soils, or to increase the level of free phosphorus in plants used
CC for animal consumption, or to decrease the level of free phosphate in
CC seeds where phosphate levels would otherwise be deleterious to
CC germination. The invention also provides means for identifying compounds
CC that bind to and/or increase or decrease the enzymatic activity of the
CC Zmpt polypeptides.
XX
SQ Sequence 554 AA;
AAY53788 Length: 554 May 30, 2002 09:25 Type: P Check: 9398 ..

251 RYTALIEGNA KOANDMQKV MDVEIQAEQD KLARYKAND YPLLREPAR
301 RHGLHLIGTA TTWFLDIAP YSQNLTKQDI FPAIKLTSPV DDINALKEVF
351 EISKAMFLVA LLGTFPGYW TVALIDKMGR YLIQLGFEM MSVFMIMGV
401 MYNDLKNKHT TLFALFYALT FEFANFGPNS TTFVLPALFL PTRVRSTCHA
451 ISAASKAGA IVAAFGVQSL TLKGDVGHK KALIIISVTN ILGFFFTFLV
501 PETMGRSLEE ISGEDGNVEN GPGAPAGVAM GVADVSKDDK MPVSTEWQS
551 SMHA
11AA_SEQUENCE 1.0
ID ABB63883 standard; Protein; 2040 AA.
XX
AC ABB63883;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 18441.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL07986.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 18441; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2040 AA;
ABB63883 Length: 2040 May 30, 2002 09:25 Type: P Check: 9003 ..

201 FAERSDTML QKRTIGDAA SGATQEKLR KIVAQAQTAP RKSSAVEEPP
251 QTPQSANLTV PGHLKAPILA TTPASPAGNI LDMQKITQT PTTTASTGED
301 ETVPETPVI QVESTESPL LDGETGEATS TLAANSSEY INSVGVRFTQ
351 QSTHDVTSI SPYGLPFIQ LFRFLIILCN PLDKQNSDSM MHTGLSLTIV
401 AFEVAADNIG KYEGLLELVK DDLICRNLSL LSSERLSIFA ADLQLCFLLE
451 ESLRGLKFKQ LEAVLRKLSE IIASDNPKTP YEMRELALDN LLQLWRIPGF
501 VTELYINYDC DLYCTDMFES LTNLSKYTL SATNAVYSTH IISMDTLISV
551 IDSIERNCAA SKNSSNNRES LPEAPATGG SRHSRHSGL EGIVIDSGNS
601 VAABEKEVENI ASFINASSHR LRLQSGEGV GITSEQLAKV KQKRRLSOG
651 TERENQPEK GIQYLQEHGI LNAELDPQV ALFLRENPGI DKMIGEYIS
701 KKKNVDSKIL INEVDSEDFI GLRVQALRL YLETERLPGE APLIFLVLEH
751 FSDHWHKQNO DPEANVDAF RLAVYAIMLN MDQHSNAKR LNVPMTLEDF
801 TKNLRLNGG EDEQEMLAQ VENAIKNEEI VMPAEQTGLV RENYQWKVLL
851 RRGDTHDGHF HYVHDASYDV EIFNIWVGAS LSALSEFMFK STETGYORTL
901 AGFSKSAIS AHYNLHSDFD ALVLTLCFT TLSSVEQHE PAPANNETOQ
951 AVNEGLNGKA QAAMRTVELL VHDYGDCLRE SWKHLLDLYL QLFRLKLPRK
1001 SLIEVEDECE ANGKAMLILE KPREKQESGL FSSLYSFSS EGOREPTYEE
1051 QDFIKLGRKC IKECQLDQML QESKEVQLES LQELLKCVLA LKAPQGHKS
1101 IGLPYAEDQT VFWMEFLVKI VVHNRDRMP LMPAVRDQMY LLMGSASCG
1151 YDYLLNRCIV AVLKLAITYLM RNEELCPIVL QSLKMLMLK PALLLRISKQ
1201 ISIGIYELLK TSAQNIHSEQ DMQIIFNLLE CVGAGAVPRN YDDAQLPLRP
1251 NSAKSDGAI SGEEDATAPV ERGYTSDSEI TKASAPAVS SPSAENWILV
1301 NNKSELTTA SRQSPSLS APPVNTLVYN COLLDHAPFA LEKWCDSLAF
1351 IYRSVAHITP YNEACVRCI RIFVEACRDG GIRQRRKLES AAKOKSSKKR
1401 SERKPGMASS ASSSNLTLT GDRSDNQING NAAEQEDLAQ RYEQLSIQLL
1451 DLMYTLYTRT AQIFRWMAEE GCIVPOSAAL WSPGWCPLAQ GIARLAMDRR
1501 REVRTHAISC LQORALLVHD LQTLSTGTEWC SCFHQVLFP LNELLPESNA
1551 AGQDALLLE ESRIRATIM SKVFLQHLTY LTELGNAFNE LMLDILDYIE
1601 RFMKVGSDDL SEQWQELKN MLVMSHVRV FHNODGSLQ ALMELTWRI
1651 GEFLPNLKEE LFHEDSITP AVSLDYQIR FAHSTGTTY SLSPLXYGS
1701 PKISSISASK CDTLPATORK KPMAKLRSL KLLKKAVALA APQOQLPAVT
1751 ILPQTQVSN ELVSAFTPP AATPLLGSV ESPRRSILQ PMADVLOQP
1801 PSFYFAQPII VPPQPAVTD PIPPSTLLPD LVNEATAAAY QATTTSPTHS
1851 PQEAGQPASI VQQTNIVTN NFNYSYALIV PMAETTAEQ FGQOQOQLLY
1901 QQYVOQYQAO QQQLPAPASD PAINVPISHL LAGNAYPSLP KMPQASIVHS
1951 FAPYVESQAA TSGAGTAAD IYQEVQNPY NLTLQOHPQ QLHQOQOQOQ

2001 QOATGMANAF PAVATPANYF NVNVDPSIS PGEELLYGQO
11AA_SEQUENCE 1.0
ID ABB69566 standard; Protein; 337 AA.
XX
AC ABB69566;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 35490.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR MPI; 2001-656860/75.
DR N-PSDB; ABL13669.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 35490; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 337 AA;
ABB69566 Length: 337 May 30, 2002 09:25 Type: P Check: 7480 ..
1 MCRISEFEKD ISIGVSVFTL TALSGERYCA IVNPLRKLQT KPLTVFTAVM
51 IWLAILLGM PSVLESDIKS YPVFTATGNN TIEVCSFPRD PEYAKFMVAG
101 KALVYLLPL SIIGALYIMM AKRLHMSARN MPEQOSMOS RTQARARLHV
151 ARMVVAFVVV FFICFFPYHV FELWYHFYPT AEEDFEDEFWN VLRIVGCTS
201 FLNSGVNPVA LYCVSGVERQ HFNRYLCCIC VKRQPHLRQH STATGMDNT
251 SVMSMRSTY VCGTAGNLRA SLHRNSNHGV GGAGGCVGGG VSGRGVSFH
301 RQDSMPLQHG NAHGGAAGG SSGIAGAGRT AAVSEKR
11AA_SEQUENCE 1.0
ID AAU45398 standard; Protein; 89 AA.
XX
AC AAU45398;
XX

DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #6294.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59525.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 6593; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 89 AA;
AAU45398 Length: 89 May 30, 2002 09:25 Type: P Check: 183 ..
1 NESVPNRYHT SGYIRGSRA TSLKPSTMT AKASIHYSYSP TTMNRNLRC
51 RSLGRCASP SSSSQRVSP AFPRHMDTLH EIRAGSYSP
11AA_SEQUENCE 1.0
ID ABG03782 standard; Protein; 146 AA.
XX
AC ABG03782;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #3773.
XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS67969.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 34141; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 146 AA;
ABG03782 Length: 146 May 30, 2002 09:25 Type: P Check: 3344 ..
1 LKKIDRLPFK DFWQRFSLDL KALAVKQORT VYRLTLVKAW NVDELQAVAO
51 IVSLGNPDFI EVKFKIGGEW WTWIDYNRFQ ELIQEYEDSG GSKTFSAKDY
101 MARTPHWALF GASERGFDPK DTRHSEERNK FKGLIFWVE IYLIFQ
11AA_SEQUENCE 1.0
ID ABG03783 standard; Protein; 244 AA.
XX
AC ABG03783;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #3774.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.

XX WO200175067-A2.
PN
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR N-PSDB; AAS67970.
DR
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity .
XX
XX Claim 20; SEQ ID No 34142; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 244 AA;
ABG03783 Length: 244 May 30, 2002 09:25 Type: P Check: 8253 ..
1 GIAGFVNIHL DLSLFLTGP GYKAERFEXR MTAHKCALSL VGEPIMYPEI
51 NRFLKLLHQC KISSFLVTNA QFPAEIRNLE PVTQLYVRVD ASTKDSLAKKI
101 DRPLFKDFWQ RFLDSVKALA VKYLQIRIGSR TPNMTKIYSY CPAVHPAEPT
151 DMKSWPSLFE VPTSLEYCPF YLQLVESADA ESTQKYRRLT AYYIPVTEP
201 PLITKEPSC L WKOAEFGDLG KHWLVLEQFS STRVOEHGVC WGVE
11AA_SEQUENCE 1.0
ID ABG05207 standard; Protein; 636 AA.
XX
XX ABG05207;
AC
XX
DT 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #5198.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.

XX WO200175067-A2.
PN
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR N-PSDB; AAS69394.
DR
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity .
XX
XX Claim 20; SEQ ID No 35566; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 636 AA;
ABG05207 Length: 636 May 30, 2002 09:25 Type: P Check: 5356 ..
1 MLISYPISE VLEVLDLCQE PRPLQALKFV SPLQNDGVKR GAENTDLWLI
51 SEOSDFKIQK VSTMRSEISA SASKALICHG TYTLRMGCMC MYLSVSPSSL
101 ENPKGNSKV VDSGDNSHIY SEGKEPSSLI RGALISRCRS KHLILFSSVF
151 LPLGPKCKQS QEVVRPIIMSH LCGWKSCKGN PREGSGHEI DLHKYNQILIL
201 GKDLMPFTRV TVHWGKGNDQ TFQGLQDTGS ELMILPGDPK RHCSPPVKVG
251 SYGGQVINGV LAGVRLTVAA VSDVSVSLLEQ INISPGTWYA AIDVANAFPS
301 IPVMAHLKQ FAFSWOGQOY TFTVLPRGW EINPTKIQGP STSVKFLGIQ
351 WYGACQDIPS KYKEKLHLA PPTTKKEAQR LVLDLFGFRQ HIPHVGVLWM
401 PIYQVTRKAA SFEWGPQOEK ALQOVQAAVQ AALPLGPYDP ADPLVLEVDA
451 VMSDQOAPIG ESQRRPLGFV SKALPPSAGN YSPFERQLFV CYWALLETEH
501 LTMGHQVTMQ PELPIMNWVP SNQSSYKVGH AQOHSIIKWK WYIRDCARAH
551 PEGTIRIHRS RNOGVEVEVV PLTITPSDPL AKALLPVPAV LCSAGTGVLV

```
601 PEGTLPFGD TTMIPLNWKL RLPFGHGLL LPLSQD
!!AA_SEQUENCE 1.0
ID ABG05563 standard; Protein; 636 AA.
XX
AC ABG05563;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #5554.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS69750.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID No 35922; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 636 AA;
ABG05563 Length: 636 May 30, 2002 09:25 Type: P Check: 5356 ..
1 MLSTYPISQE VLEVLDLQOE PRPLQALKFV SPLQNDGVKR GAENTDMLLI
51 SEQSDFKIQK VSTMRESIA SASKALICHG TYTLRMGCMC MYLSTVSPSSL
101 ENPKGNSKVV VDSGDNSHIY SEGKEPSSLI RGALISRCRS KHLLEFSSVF
151 LPLGPKCKQS QEVRIILMSH LCGWKSCKGN PREGSHGEI DLHKYNQLIL
```

```
201 GKDLMPFTRV TVHWGKGNDQ TFOGLQDTGS ELMILPGDPK RHCSPPYKVG
251 SYGGQVINGV LAQVRLTVAA VSDVSLLEQ INISPGTWXA AIDVANAFSS
301 IPVMAHLKQ FAFSWOGQOY TFTVLPRGW EINPTKIQCP STSVKFLGICQ
351 WYGACQDIPS KYKEKLLHLA PPTTKKEAQR LVDLEGFRRQ HIPHVGULLM
401 PIYQVTRKAA SFEMGPEQEK ALQVQAAYQ AALPLGPYDP ADPLVLEVDA
451 VMSPOQAPIG ESQRRPLGEW SKALPPSAGN YSPFERQLEF CYWALLETEH
501 LTMGHQVTMQ PELPIMNWP SNQSSYKVGH AQOHSIITWK WYIRDCARAH
551 PEGTIRIHRS RNQVEVEVY PLTITPSDPL AKALLPVPAAT LCSAGTGVLY
601 PEGTLPFGD TTMIPLNWKL RLPFGHGLL LPLSQD
!!AA_SEQUENCE 1.0
ID ABG06670 standard; Protein; 236 AA.
XX
AC ABG06670;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6661.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS70857.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID No 37029; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
```

CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 236 AA;

ABG0670 Length: 236 May 30, 2002 09:25 Type: P Check: 936 ..

1 LSTQAKRRK EPWSLRELRL GSSPGGCDL LFGVLWLTFF SSWAPLRSP
51 VPASEAACSA PGLAAMLRA SACAGTWSGR SRCSSHTDV HTVASLKLKLY
101 LRELPEPVIP YAKYEDFLSC AKLSKEEA GYKELAKQVK SLPVVNNLL
151 KYICRFLEDEV QSYSGVNKMS VONLATVFGP NILRPKVEDP LTIMEDILKI
201 KEDKIERKGC EKEHRNDLND FKSLSITEV AFYPLD

!!AA_SEQUENCE 1.0
ID ABG06767 standard; Protein: 636 AA.

AC ABG06767;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #6758.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS70954.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 37126; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 636 AA;

ABG06767 Length: 636 May 30, 2002 09:25 Type: P Check: 5356 ..

1 MLSTYPIQOE VLEVLDLCQE PRPLQALKFV SPLQNDGVKR GAENTDMLLI
51 SEQSDFKIQK VSTMRSERIA SASKALICHG TYTLRMGCMC MYLSVSPSSL
101 ENPKGNSKVV VDSGDNSHIY SEGKEPSSLI RGALISRCRS KHLLEFSSVF
151 LPLGPKCKQS QEVPRILMSH LCGWKSCKGN PREGSGHEI DLHKYNQLIL
201 GKDLMPFTRV TVHMGKNDQ TFOGLQDTGS ELMILIPGDPK RHCSPPVKVG
251 SYGQVINGV LAQVRLTVAA VSDVVSLEEQ INISPGTWYA AIDVANAFSS
301 IPVHMAHLKQ FAFSWOGQOY TFTVLPQRGW EINPTKIQGP STSVKFLIGIQ
351 WYGACODIPS KVKERLHLA PPTTKKEAQR LVDLGFRRQ HIPHVGVLWM
401 PIYQVTRKAA SFEMGPEQEK ALQOVQAAVQ AALPLGPYDP ADPLVLEVDA
451 VMSPOQAPIG ESQRRPLGFV SKALPPSAGN YSPPERQLFV CYWALLETEH
501 LTMGHQVTMQ PELPIMNWVP SNQSSYKVGH AQQHSIIKKW WYIRDCARAH
551 PEGTIRIHRN RNQGVVEVEV PLTITPSDPL AKALLPVPAT LCSAGTGVLV
601 PEGTLPPLPGD TTMIPLNMWL RLPPGHFGLL LPLSQD

!!AA_SEQUENCE 1.0
ID ABG17402 standard; Protein: 636 AA.

AC ABG17402;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #17393.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS81589.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 47761; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 636 AA;

ABG17402 Length: 636 May 30, 2002 09:25 Type: P Check: 5356 ..

1 MLSTYPISQE VLEVLDLQCE PRPLQALKFV SPLQNDGVKR GAENTDMLLI
51 SEQSDFKIQK VSTMRSSEIA SASKALICHG TYTLRMGCMC MYLSVSPSSL
101 ENPKGNSKVV VDSGDNSTHY SEGKEPSSLI RGALISRCRS KHLLEFSVF
151 LPLGPKCKOS QEVRPILMSH LCGWKSCKGN PREGSHGEI DLHKYNQULI
201 GKDLMPFTRV TVHWGKGNDQ TFOGLQDTGS ELMILPGDPK RHCSPPYKVG
251 SYGQVINGV LAQVRLTVAA VSDVVSILEQ INISPGTWYA AIDVANAFES
301 IPVMAHLKQ FAFSWGGOY TFTVLPQRGW EINPTKIQGP STSVKFLGIQ
351 WYGACQDIPS KVEKELHLA PPTTKKEAQR LVDLFGFRQ HIPHVGVLWM
401 PIYQVTRKAA SFEMGPEDCK ALQOVQAAVQ AALPLGPYDP ADPLVLEVDA
451 VMSPOQAPIG ESQRRPLGFV SKALPPSAGN YSPFERQLFV CYWALLETEH
501 LTMGHQVTMQ PELPIMNVP SNOSYKVGH AQHSHIKWK WYIRDCARAH
551 PEGTIRIHRN RNOGVEVEVY PLTTTPSDPL AKALLPVPAAT LCSAGTGVLY
601 PEGGTLPPGD TTMIPLNWKL RLPGRHFGLL LPLSQD

11AA_SEQUENCE 1.0

ID ABG29764 standard; Protein; 198 AA.

XX AC ABG29764;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #29755.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217;

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS93951.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity

XX PS Claim 20; SEQ ID No 60123; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 198 AA;

ABG29764 Length: 198 May 30, 2002 09:25 Type: P Check: 9364 ..

1 VRAVFSLREQ IRSHEIRRSA AVSNHQHFR ACGHIDRRTV QTLAHLTRF
51 GDKGVTRPED FVHFWRHFRF KGDPDLTFWP SAPVTATILR IPLANGFRFH
101 DHKSSCMTRV PQVPEDPVKS HAPQOFWINH GKFAWYDVA IGIPCPAGR
151 SCVILLGDV REVTRVSNV PAKRYIGGSP RRGFINSQMA PGIIISPA

11AA_SEQUENCE 1.0

ID ABB29544 standard; Peptide; 27 AA.

XX AC ABB29544;

XX DT 01-FEB-2002 (first entry)

XX DE Peptide #2195 encoded by breast cell single exon nucleic acid probe.

XX KW Human; microarray; single exon probe; gene expression; breast;

XX KW disease; cancer.

XX OS Homo sapiens.

XX PN WO200157271-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00662.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.


```

PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-496933/54.
DR
XX
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
XX Claim 27; SEQ ID NO 12512; 327pp + sequence listing; English.
PS
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 27 AA;
SQ
ABB29544 Length: 27 May 30, 2002 09:25 Type: P Check: 8372 ..

1 SLKKIDRPLF KDFWQGFIDS LKALAVK
11AA_SEQUENCE 1.0
ID ABB34423 standard; Peptide; 113 AA.
XX
XX ABB34423;
AC
XX
XX 04-FEB-2002 (first entry)
DT
XX
XX Peptide #1929 encoded by human foetal liver single exon probe.
DE
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200157277-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00669.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX
XX 26-MAY-2000; 2000US-0207456.
PR
XX
XX 30-JUN-2000; 2000US-0608408.
PR
XX
XX 03-AUG-2000; 2000US-0632366.
PR
XX
XX 21-SEP-2000; 2000US-0234687.
PR
XX
XX 27-SEP-2000; 2000US-0236359.
PR
XX
XX 04-OCT-2000; 2000GB-0024263.
PA
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX

```

```

DR WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
PT
XX
XX Claim 27; SEQ ID NO 27058; 639pp + sequence listing; English.
PS
XX
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 113 AA;
SQ
ABB34423 Length: 113 May 30, 2002 09:25 Type: P Check: 1841 ..

1 FITSLPDM EATLWIGLRWT AYEKINKWTD NRELTYSNFH PLLVSGRLRI
51 PENVSFFKNQ YSSYLCEILE SFIHFWTKQN PTVLFHLYEM SRIGTSKIYS
101 RLVYARGYGK WKD
11AA_SEQUENCE 1.0
ID ABB19835 standard; Protein; 113 AA.
XX
XX ABB19835;
AC
XX
XX 23-JAN-2002 (first entry)
DT
XX
XX Protein #1834 encoded by probe for measuring heart cell gene expression.
DE
XX
XX Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200157274-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00666.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX
XX 26-MAY-2000; 2000US-0207456.
PR
XX
XX 30-JUN-2000; 2000US-0608408.
PR
XX
XX 03-AUG-2000; 2000US-0632366.
PR
XX
XX 21-SEP-2000; 2000US-0234687.
PR
XX
XX 27-SEP-2000; 2000US-0236359.
PR
XX
XX 04-OCT-2000; 2000GB-0024263.
PA
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX
XX WPI; 2001-48899/53.
DR
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
PT
XX
XX Claim 15; SEQ ID NO 21605; 530pp; English.
PS
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,

```

CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 113 AA;

ABBI9835 Length: 113 May 30, 2002 09:25 Type: P Check: 1841 ..

1 FITSLPDM EATLMIGLRWT AY EKINKWTD NRELTYSN EHPILVSGRLRI
51 PENVSFFKNQ YSSYLCEILE SFIHFWTKON PTVLFHLYEM SRIGTSKIVS
101 RLVIAAGYGC WKD

11AA_SEQUENCE 1.0
ID AAB86968 standard; Protein: 506 AA.

XX AAB86968;

XX 11-DEC-2001 (first entry)

DE D. melanogaster peptide receptor SEQ ID 34.

KW Insect; fruitfly; peptide receptor; plant protection; insecticide.

OS Drosophila melanogaster.

PN DE10013618-A1.

PD 20-SEP-2001.

PF 18-MAR-2000; 2000DE-1013618.

PR 18-MAR-2000; 2000DE-1013618.

PA (FARB) BAYER AG.

PI Antonicek H, Friedrich G, Schulte T;

DR WPI; 2001-571695/65.

DR N-PSDB; AAH49421.

XX New polypeptides from Drosophila melanogaster have biological activity
PT of peptide receptor, useful to find new compounds for plant protection
PT and insecticides -

PS Claim 2; Page 90-92; 128pp; German.

XX This invention describes novel polypeptides (P1) from Drosophila
CC melanogaster having the biological activity of a peptide receptor.
CC Molecules of the invention are used to find new plant protection
CC compounds or insecticides, or to find genes encoding a polypeptide
CC involved in the structure of functionally similar receptors in insects
CC This sequence represents a Drosophila melanogaster (fruitfly) peptide
CC receptor described in the method of the invention.

XX Sequence 506 AA;

AAB86968 Length: 506 May 30, 2002 09:25 Type: P Check: 4326 ..

1 MYASIMDVGO TLAARLADSD GNGANDSGLL ATGGGLEQEO EGLALDMGHN
51 ASADGIYVY VPVLDREPTY IYTVLYTLIF IYGVLCNGTL VIIFRHRSM
101 RNIPNTYILS LALADLVIL VCVPVATIVY TQESWPFERN MCRISEFFKD
151 ISIGVSVFTL TALSGERYCA IVNPLRKLOT KPLIVFTAVM IWIILLLGM
201 PSVLFSDIKS YPVFTATGNN TIEVCSPPRD PEYAKFWVAG KALVYLLPL

251 SIIGALYIM AKRLHMSARN MPEGOQSMQS RTQARARLHV ARMVAVFVV
301 FFICFPYHV FELWYHFYPT AEEDFDEFWN VLRIYGFCTS AVKLKNVSA
351 IWWLNSLPLV VKAICILAKW LPKLVROPFG LYCVSGVERQ HFNRYLCCTC
401 VKROPHLRQH STATGMDNT SVMSMRSTY VGTAGNLRA SLHRNSNHGV
451 GGAGGVGGG VSGSRVGSFH RQDSMPLQHG NAHGAGCGG SSGLAGGRT
501 AAVSEK

11AA_SEQUENCE 1.0
ID AAE11926 standard; Protein: 288 AA.

XX AAE11926;

DT 18-DEC-2001 (first entry)

DE Human CG27 (or C869) receptor protein #1.

KW Human; apolipoprotein; lipase; lipoprotein receptor; ALIr; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipaeitic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective.

XX Homo sapiens.

PN WO200179446-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US12529.

PR 14-APR-2000; 2000US-197137P.

PR 20-JUN-2000; 2000US-0598042.

PR 03-AUG-2000; 2000US-0631451.

PR 22-SEP-2000; 2000US-0667298.

PR 17-NOV-2000; 2000US-0714936.

XX (HYSE-) HYSEQ INC.

XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;

DR WPI; 2001-611724/70.

DR N-PSDB; AAD19221.

XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
PS Claim 10; Fig 4; 266pp; English.

XX The invention relates to polynucleotides encoding proteins CG122, CG179,
CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALIr)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of these sequences. ALIr polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALIr proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC Vectors comprising these DNA and protein sequences are also useful for
CC producing ALIr proteins. The nucleic acids and polypeptides of the

CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human CG27 (or C869)
CC receptor protein.

XX Sequence 288 AA;

AAE11926 Length: 288 May 30, 2002 09:25 Type: P Check: 2953 ..

1 MSLHSQASAT TRPEPRRTE HRAPSSWRP VALTLTLCL VLLIGLALG
51 LLEFQYQLS NTGDTISQM EERLGNSTQE IQLQVQNIK LAGSLQHVAE
101 KLCRELYNKA GGYTRNMVPA SASSESLROL PHMGESAAPH RCSPCTEQWK
151 WHGDCYQFY KDSKSWEDCK YFCISENSTM LKINKQEDLE FAASQSYSEF
201 FYSYWTGLLR PDSGKAWLWM DGTPTPTSELF HIIDVTSPP SRDCVAAILG
251 MIFSKDCKEL KRCVCERRAG MKKPESLHVP PETLGECD

11AA_SEQUENCE 1.0
ID AAE11932 standard; Protein; 280 AA.

XX AAE11932;

DT 18-DEC-2001 (first entry)

DE Human CG27 (or C869) receptor protein variant #2.

XX Human; apolipoprotein; lipase; lipoprotein receptor; ALLr; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective; variant.

XX Homo sapiens.

XX Key Location/Qualifiers
FT Domain 148..166
FT /note="C-type lectin domain"

XX WO200179446-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US12529.

XX 14-APR-2000; 2000US-197137P.
PR 20-JUN-2000; 2000US-0598042.
PR 03-AUG-2000; 2000US-0631451.
PR 22-SEP-2000; 2000US-0667298.
PR 17-NOV-2000; 2000US-0714936.

XX (HYSE-) HYSEQ INC.

XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;

XX WPI; 2001-611724/70.
DR N-PSDB; AAD19227.

XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX Claim 10; Fig 4; 266pp; English.

XX The invention relates to polynucleotides encoding proteins CG122, CG179,

CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALLr)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of these sequences. ALLr polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALLr proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC Vectors comprising these DNA and protein sequences are also useful for
CC producing ALLr proteins. The nucleic acids and polypeptides of the
CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human CG27 (or C869)
CC receptor protein variant.

XX Sequence 280 AA;

AAE11932 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..

1 MQAKYSTRD MLDDGDPTM SLHSQASATP RHPEPRTEH RAPSSWRPV
51 ALTLTLCLV LLIGLALGL LFFQYQLSN TGDTISQME ERLGNSTQEL
101 QSLQVQNIK LAGSLQHVAEK LCRELYNKAAG AHRSPCTEQ WKWHGDCNYQ
151 FYKDSKSWED CKYFCLSENS TMLKINKQED LEFAASQSYS EFFYSYWTGL
201 LRPDSGKAWL WMDGTPTTSE LFHIIDVTS PRSRDCVAAIL NGMIFSKDCK
251 ELKRCVCERR AGMKPESLH VPPETLGECD

11AA_SEQUENCE 1.0
ID AAE11933 standard; Protein; 314 AA.

XX AAE11933;

DT 18-DEC-2001 (first entry)

DE Human CG27 (or C869) receptor protein variant #3.

XX Human; apolipoprotein; lipase; lipoprotein receptor; ALLr; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective; variant.

XX Homo sapiens.

XX Key Location/Qualifiers
FT Domain 190..284
FT /note="C-type lectin domain"

XX WO200179446-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US12529.

XX 14-APR-2000; 2000US-197137P.
PR 20-JUN-2000; 2000US-0598042.
PR 03-AUG-2000; 2000US-0631451.
PR 22-SEP-2000; 2000US-0667298.
PR 17-NOV-2000; 2000US-0714936.

XX (HYSE-) HYSEQ INC.

XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX
DR WPI: 2001-611724/70.
DR N-PSDB: AAD19228.
XX
PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX
XX Claim 10; Fig 4; 266pp; English.
PS
CC The invention relates to polynucleotides encoding proteins CG122, CG179,
CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALLr)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of these sequences. ALLr polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALLr proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC vectors comprising these DNA and protein sequences are also useful for
CC producing ALLr proteins. The nucleic acids and polypeptides of the
CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human CG27 (or C869)
CC receptor protein variant.
XX
XX Sequence 314 AA:
SQ
AAE11933 Length: 314 May 30, 2002 09:25 Type: P Check: 6095 ..
1 MQAKYSTMD MDDGDGDTM SLHSQASAT RHPEPRTEH RAPSTWRPV
51 ALTLTLTCLV LLIGLALGL LFFQYYQLSN TGQDTISQME ERLGNTSQEL
101 QSLQVONIKL AGSLQHVAEK LCRELKYNKAG GYTRNMVPAS ASSESLROLR
151 HMGESAAHR CSPCTEQWKW HGDNCYQFYK DSKSWEDCKY FCLSENSTML
201 KINKQEDLEF AASQSYSEFF YSYWTGLLRP DSGKAWLMMD GTPFTSELFH
251 IIVDTSPRS RDCVALNGM IFSKDKELK RCVCERRAGM VKPESLHVPP
301 ETLGEGDMHH HHHH
11AA_SEQUENCE 1.0
ID AAE11934 standard; Protein; 247 AA.
XX
AC AAE11934;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human CG27 (or C869) receptor protein variant #4.
XX
KW Human; apolipoprotein; lipase; lipoprotein receptor; ALLr; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective; variant.
XX
OS Homo sapiens.
XX

FH Key Location/Qualifiers
FT Domain 130..224
FT /note- "C-type lectin domain"
XX
XX WO200179446-A2.
XX
XX
PD 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US12529.
XX
XX 14-APR-2000; 2000US-197137P.
PR 20-JUN-2000; 2000US-0598042.
PR 03-AUG-2000; 2000US-0631451.
PR 22-SEP-2000; 2000US-0667298.
PR 17-NOV-2000; 2000US-0714936.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX
DR WPI: 2001-611724/70.
DR N-PSDB: AAD19229.
XX
PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX
XX Claim 10; Fig 4; 266pp; English.
PS
XX The invention relates to polynucleotides encoding proteins CG122, CG179,
XX CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
XX involved in lipid metabolism and cardiovascular disease such as human
XX apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
XX and protein sequences are useful for treating or preventing disorders
XX associated with apolipoproteins, lipases and lipoprotein receptor (ALLr)
XX expression and for treating lipid metabolism, cardiovascular diseases
XX and thrombosis. Antibodies against these proteins are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of these sequences. ALLr polypeptides are also
XX useful for identifying agents (agonists and antagonists) that bind to
XX them and cells expressing ALLr proteins are useful for identifying a
XX therapeutic agent for use in treatment of a pathology related to
XX aberrant expression or physiological interactions of this polypeptide.
XX vectors comprising these DNA and protein sequences are also useful for
XX producing ALLr proteins. The nucleic acids and polypeptides of the
XX invention are also useful for the treatment of occlusive cardiovascular
XX diseases, myocardial infarction, cerebral ischaemia, angina, arterial
XX thrombosis, coronary artery thrombosis and cerebral artery thrombosis
XX or intracardiac thrombosis and stroke. The nucleotides of the invention
XX are used in gene therapy. The present sequence is human CG27 (or C869)
XX receptor protein variant.
XX
XX Sequence 247 AA:
SQ
AAE11934 Length: 247 May 30, 2002 09:25 Type: P Check: 89 ..
1 MQAKYSTRD MDDGDGDTM SLHSQASAT RHPEPRTVF QYYQLSNTGQ
51 DTISQMEERL GNTSQELQSL QVONIKLAGS LQHVAEKLKR ELYNKAGAH
101 CSPCTEQWKW HGDNCYQFYK DSKSWEDCKY FCLSENSTML KINKQEDLEF
151 AASQSYSEFF YSYWTGLLRP DSGKAWLMMD GTPFTSELFH IIVDTSPRS
201 RDCVALNGM IFSKDKELK RCVCERRAGM VKPESLHVPP ETLGEGD
11AA_SEQUENCE 1.0
ID AAE11938 standard; Protein; 280 AA.
XX
AC AAE11938;
XX

DT 18-DEC-2001 (first entry)
XX
DE Human lipid metabolism related protein #1.
XX
KW Human; apolipoprotein; lipase; lipoprotein receptor; ALlr; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipaeamic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective.
XX
OS Homo sapiens.
XX
PN WO200179446-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US12529.
XX
PR 14-APR-2000; 2000US-197137P.
PR 20-JUN-2000; 2000US-0598042.
PR 03-AUG-2000; 2000US-0631451.
PR 22-SEP-2000; 2000US-0667298.
PR 17-NOV-2000; 2000US-0714936.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX
DR WPI; 2001-611724/70.
DR N-PSDB; AAD19233.
XX
PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX
PS Claim 10; Page 252; 266pp; English.
XX
CC The invention relates to polynucleotides encoding proteins CG122, CG179,
CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALlr)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of these sequences. ALlr polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALlr proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC Vectors comprising these DNA and protein sequences are also useful for
CC producing ALlr proteins. The nucleic acids and polypeptides of the
CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human protein
CC related to proteins involved in lipid metabolism.
XX
SQ Sequence 280 AA:

AAE11938 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..

1 MQAKVSTRD MLDDGDTM SLHSQASAT RHPEPRTEH RAPSSTWRPV
51 ALTLTLTCLV LLIGLALGL LFFQYQLSN TGQDTISQME ERLGNTSQEL
101 QSLQVONIKL AGSLQHVAEK LCRELYNKAG AHRCSPCTEQ WKWHGDNQCYQ

151 FYKDSKSWED CKYFCLSENS TMLKINKQED LEPASQSYS EFFYSWTGL
201 LRPDSGKAWL WMDGTPFTSE LFHIIIDVTS PSRDCVAIL NGMIFSKDK
251 ELKRCVCERR AGMWKPESLH VPPELTGEGD
11AA_SEQUENCE 1.0
ID AAE11939 standard; Protein; 307 AA.
XX
AC AAE11939;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human lipid metabolism related protein #2.
XX
KW Human; apolipoprotein; lipase; lipoprotein receptor; ALlr; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipaeamic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective.
XX
OS Homo sapiens.
XX
PN WO200179446-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US12529.
XX
PR 14-APR-2000; 2000US-197137P.
PR 20-JUN-2000; 2000US-0598042.
PR 03-AUG-2000; 2000US-0631451.
PR 22-SEP-2000; 2000US-0667298.
PR 17-NOV-2000; 2000US-0714936.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX
DR WPI; 2001-611724/70.
DR N-PSDB; AAD19234.
XX
PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX
PS Claim 10; Page 254-255; 266pp; English.
XX
CC The invention relates to polynucleotides encoding proteins CG122, CG179,
CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALlr)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of these sequences. ALlr polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALlr proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC Vectors comprising these DNA and protein sequences are also useful for
CC producing ALlr proteins. The nucleic acids and polypeptides of the
CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human protein
CC related to proteins involved in lipid metabolism.

XX Sequence 307 AA;
SQ
AAE11939 Length: 307 May 30, 2002 09:25 Type: P Check: 2921 ..
1 MQAKYSTTRD MDDDDGDTTM SLHSQASATP RHEPRRTVE QYQLSNTGQ
51 ALTLTLTCLV LLIGLALGL LFFQYQLSN TGQDTISQME ERLGNTSQEL
101 QSLQVONIKL AGSLQHVAEK LCRELKYNKAG GYTRNMVPAS ASSESLRQLP
151 HMGESAAAHK CSPCTEQWKW HGDNCYQFYK DSKSWEDCKY FCLSENSTML
201 KINKQEDLEF AASQSYSEFF YSYWTGLLRP DSGKAWLMMD GTPETSELPH
251 IIDVTSPRS RDCVAILNGM IFSKDCKEK RVCERRAGM VKPESLHVP
301 ETLGEGD
11AA_SEQUENCE 1.0
ID AAE11943 standard; Protein; 281 AA.
XX
AC AAE11943;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human CG27 (or C869) receptor protein variant #5.
XX
KW Human; apolipoprotein; lipase; lipoprotein receptor; ALlr; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective; variant.
XX
OS Homo sapiens.
XX
PN WO200179446-A2.
XX
PD 25-OCT-2001.
XX
PE 16-APR-2001; 2001WO-US12529.
XX
PR 14-APR-2000; 2000US-197137P.
PR 20-JUN-2000; 2000US-0598042.
PR 03-AUG-2000; 2000US-0631451.
PR 22-SEP-2000; 2000US-0667298.
PR 17-NOV-2000; 2000US-0714936.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX
DR WPI; 2001-611724/70.
DR N-PSDB; AAD19238.
XX
PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX
PS Claim 10; Page 265-266; 266pp; English.
XX
CC The invention relates to polynucleotides encoding proteins CG122, CG179,
CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALlr)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated

CC with altered levels of these sequences. ALlr polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALlr proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC Vectors comprising these DNA and protein sequences are also useful for
CC producing ALlr proteins. The nucleic acids and polypeptides of the
CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human CG27 (or C869)
CC receptor protein variant.
XX
SQ Sequence 281 AA;
AAE11943 Length: 281 May 30, 2002 09:25 Type: P Check: 7726 ..
1 MQAKYSTTRD MDDDDGDTTM SLHSQASATP RHEPRRTVE QYQLSNTGQ
51 DTISQMEERL GNTSQELQSL QVONIKLAGS LQHVAEKLKR ELYNKAGYTT
101 RNMYVPASASS ESLRQLPHMG ESAAAHRCSP CTEQWKWHD NCYQFYKDSK
151 SWEDCKYFCL SENSTMLKIN KQEDLEFAAS QSYSEFFYSY WTGLLRPDSG
201 KAWLMMDGTP FTSELPHIII DVTSPRSRDC VAILNGMIFS KDCKELKRCV
251 CERRAGMVKP ESLHVPETL GEGDMHHHH H
11AA_SEQUENCE 1.0
ID AAU20199 standard; Protein; 85 AA.
XX
AC AAU20199;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human novel endocrine antigen, SEQ ID No 256.
XX
KW Human; endocrine antigen; cytostatic; antifertility; antidiabetic;
KW thyroid-active; adrenal-active; androgenic; gastric; gene therapy;
KW antisense-therapy; antibody; endocrine disorder; hormone imbalance;
KW reproductive disorder; endocrine cancer; pancreatic disorder;
KW diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder;
KW hyperthyroidism; hypothalamic disorder; vanishing testes syndrome.
XX
OS Homo sapiens.
XX
PN WO200155319-A2.
XX
PD 02-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01335.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX MPI; 2001-457726/49.
XX DR N-PSDB; AAS32302.
XX
PT Isolated polypeptide for treating, preventing and prognosing disorders
PT related to the endocrine system including endocrine disorders,
PT reproductive disorders, and gastrointestinal disorders and also for
PT testing and detection e.g. diagnosis -
XX
XX
PS Claim 11; SEQ ID No 256; 558pp; English.
XX

The invention relates to cDNAs encoding novel human endocrine antigens or a fragment having biological activity, a domain, an epitope, full length protein, variant, allelic variant or a species homologue of the cDNA/antigen. The DNAs and polypeptides are useful for preventing, treating or ameliorating a medical condition when administered (e.g. by gene therapy or antisense-therapy). Identifying mutations in the genes coding for the antigens is useful for diagnosing a pathological condition or a susceptibility to a pathological condition. The DNAs, antigens and antibodies raised against the antigens useful for treating, preventing and/or prognosing disorders related to the endocrine system or hormone imbalance or reproductive disorders, cancers of endocrine tissues, disorders of the pancreas (e.g. diabetes mellitus), the adrenal glands (e.g. hirsutism), ovaries, the thyroid (e.g. hyperthyroidism), the hypothalamus and testes (e.g. vanishing testes syndrome), many examples

CC of diseases and disorders are given in the specification. The
CC present sequence represents an endocrine antigen of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 85 AA;
AAU20199 Length: 85 May 30, 2002 09:25 Type: P Check: 7019 ..
1 GCLVAQNNCQ SSSLHLILR ESQRPALARC LLEVFGSLT TLLFISHMPE
51 QSEMAIPCCR GGMEMLSRW MRECYRHKGS GGLRQ
!!AA_SEQUENCE 1.0
ID AAU27511 standard; Protein, 189 AA.
XX
AC AAU27511;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human G-Protein Coupled Receptor (GPCR) polypeptide #39.
XX
KW Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia;
KW attention deficit disorder; anxiety; depression; bipolar disorder;
KW neurological disorder; Huntington's disease; dementia; obesity; anorexia;
KW metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;
KW type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;
KW cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;
KW viral infection; immunostimulant; neuroleptic; nootropic; tranquiliser;
KW antidepressant; anorectic; gene therapy; antiviral; cardiant; alopecia;
KW neuroprotective; cytostatic; antiparkinsonian; antidiabetic; psoriasis;
KW hyperproliferative disorder; hormonal disorder; inflammatory disorder;
KW Crohn's disease.
XX
OS Homo sapiens.
XX
PN WO200162798-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-US05985.
XX
PR 25-FEB-2000; 2000US-0184939.
PR 25-FEB-2000; 2000US-0184940.
PR 25-FEB-2000; 2000US-0184941.
PR 25-FEB-2000; 2000US-0184942.
PR 25-FEB-2000; 2000US-0185042.
PR 25-FEB-2000; 2000US-0185044.
PR 25-FEB-2000; 2000US-0185208.
PR 13-MAR-2000; 2000US-0184843.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Lind P, Parodi LA, Vogell G, Wood LS;
XX
DR WPI: 2001-582042/65.
DR N-PSDB; AAS43454.
XX
PT Isolated nucleic acid molecules encoding G protein-coupled receptors
PT termed nGPCR-x, useful in the treatment and diagnosis of viral
PT infections, cancers and mental disorders (e.g. Parkinson's disease and
PT schizophrenia) -
XX
PS Claim 31; Page 76; 211pp; English.
XX
CC Sequences AAU27473-AAU27549 represent human G-protein coupled receptor
CC (GPCR) polypeptides of the invention. The proteins and their associated
CC DNA sequences can be used to identify compounds which bind to GPCR
CC polypeptides and in screening for compounds that modulate GPCR activity.
CC By screening a human subject for the presence of mutations in GPCR DNA, a
CC GPCR-related disorder or a genetic predisposition can be diagnosed. The

CC sequences can also be used for treatment and prevention of mental
CC disorders such as schizophrenia, attention deficit disorder, anxiety,
CC depression, dementia and bipolar disorder, neurological disorders such as
CC Huntington's disease, Parkinson's disease and Tourette's syndrome,
CC metabolic disorders such as obesity, anorexia and type 2 diabetes,
CC cardiovascular disorders such as thrombosis, myocardial infarction,
CC cardiomyopathy and atherosclerosis, hyperproliferative disorders such as
CC psoriasis, hormonal disorders such as alopecia, inflammatory disorders
CC such as Crohn's disease, viral infections caused by HIV and cancers.
XX
SQ Sequence 189 AA;
AAU27511 Length: 189 May 30, 2002 09:25 Type: P Check: 6166 ..
1 EVAMSRGWGE QTQLDNIMGL VFYIKVKTRK MSQRPILRCR VSIDYKLILQ
51 KAMKSLVEIE IQATHVRNCR APKYFRNVLS NKSIIKCSQP DITFGLNCFK
101 TDKKCPPLHK IITAKPVHHL TGGECKRWYT ETHRGFANDW DTKLPGRIPS
151 RFVVTKYQQ NSYVRVLCFM SYWFPQFFPG GHDKGVTEL
!!AA_SEQUENCE 1.0
ID AAU29324 standard; Protein, 280 AA.
XX
AC AAU29324;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #301.
XX
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200168848-A2.
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06520.
XX
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000MO-US14042.
PR 30-MAY-2000; 2000MO-US14941.
PR 02-JUN-2000; 2000MO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000MO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000MO-US23328.
PR 08-NOV-2000; 2000MO-US30952.
PR 01-DEC-2000; 2000MO-US32678.
PR 20-DEC-2000; 2000MO-US34956.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2001-602746/68.
DR N-PSDB; AAS46225.
XX
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX
XX Claim 11; Fig 602; 774pp; English.
PS
XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX
SQ Sequence 280 AA;
AAU29324 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..
1 MQAKYSTTRD MLDDGDTM SLHSQASAT RHPEPRTEH RAPSSTWRPV
51 ALTLTLCLV LLIGLALGL LFROYQLSN TGODTISQME ERLGNTSQEL
101 QSLQVONIKL AGSLQHVAK LCRRLYNKAG AHRCSPECTEQ WKWHGDNCTQ
151 FYKDSKSWED CKYFCLSENS TMLKINKQED LEFAASQSYS EFFYSYWTGL
201 LRPPSGKAWL WMDGTPFTSE LFHIIIDVTS PRSRDCVAIL NGMIFSKDCK
251 ELKRCVGCERR AGWVKPESLH VPPETLGEED
11AA_SEQUENCE 1.0
ID AAU32368 standard; Protein; 273 AA.
XX
XX AAU32368;
AC
XX 18-DEC-2001 (first entry)
DT
XX
DE Novel human secreted protein #2859.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX

PN WO200179449-A2.
XX
XX 25-OCT-2001.
PD
XX
XX 16-APR-2001; 2001WO-US08656.
PF
XX 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Driamanac RT;
PI
XX WPI; 2001-611725/70.
DR
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
PT
XX
XX Claim 20; Page 593-594; 765pp; English.
PS
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 273 AA;
AAU32368 Length: 273 May 30, 2002 09:25 Type: P Check: 4809 ..
1 VKGVPQVKA EREFKMTAKH CALSLVGEPI MYPEINRFLK LKHCKISSF
51 LVTNAQPPAE IRNLEPVTL YVRVDASTKD SLKKIDRPLF KDEWQRFIDS
101 VKALAVKQOR TVYRLTLVKA WNVESLQAYA QLGSLGNPDF IEVKGVTYCR
151 ESSASSLTMA HVPWHEVVO FVRELVDLIP EYEIACEHEH SNCLLIARHK
201 FKIGGEWWTW IDYNRQELI QEYEDSGGSK TFSAKDYMAR TPHWALFGAS
251 ERGFDPKDR HQRKNSKAI SGC
11AA_SEQUENCE 1.0
ID AAM55526 standard; Protein; 27 AA.
XX
XX AAM55526;
AC
XX 05-NOV-2001 (first entry)
DT
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27631.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
XX Homo sapiens.
OS
XX WO200157275-A2.
PN
XX 09-AUG-2001.
PD
XX

PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 27631; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 27 AA;
AAM55526 Length: 27 May 30, 2002 09:25 Type: P Check: 8372 ..

1 SLKKIDRPLF KDFWQFLDS LKALAVK
!!AA_SEQUENCE 1.0
ID AAM67910 standard; Protein; 27 AA.
XX
AC AAM67910;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 28216.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PI WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 28216; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 27 AA;
AAM67910 Length: 27 May 30, 2002 09:25 Type: P Check: 8372 ..

1 SLKKIDRPLF KDFWQFLDS LKALAVK
!!AA_SEQUENCE 1.0
ID AAM79021 standard; Protein; 377 AA.
XX
AC AAM79021;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1683.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PI WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK52154.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 4030-4031; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX Sequence 377 AA;
SQ AAM79021 Length: 377 May 30, 2002 09:25 Type: P Check: 5679 ..
1 MPPPPGPIG DCLRWDLDQ QDFONIQETH RLYRLKLEEL TKIQNNCTSS
51 ITROKKRLQE LALALKCKP SLPAEAGAA QELEQMKER QGLFEDMEAY
101 LPRKNGLYLS LVIGNVNVTL LSKQAKFAYK DEYEKFLYL TITILISFT
151 CRFLNSRVT DAAFNELLVW YYCTLTITRES ILINNGSRIK GWWVEHHYVS
201 TFLSGWMLTW PDGLMYQKFR NQFLSFSMYQ SFVQFLQYXY QSGCLYRLRA
251 LGERHTMDLT VEGFQSWMR GLTFLLPLFL FGHFWQLFNA LTLFNLAQDP
301 QCKEQVLMC GFPELLLFGL NFFTTLRVVH QLHHAEEAA PGAGPSPEEM
351 QTLPPSRGRG GRTGAGEPDE RAPRPLL
11AA_SEQUENCE 1.0
ID AAM80005 standard; Protein; 406 AA.
XX
AC AAM80005;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3651.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dimanac RT, Asundi V, Zhou P, Xu C, Gao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK53138.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 407; 6221p; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 406 AA;
AAM80005 Length: 406 May 30, 2002 09:25 Type: P Check: 7978 ..
1 RSMAYVKKCK NMCPRNGLH DGPEPCWLHH AAGTVSAVQA RGLQPSQSRS
51 RPRVPGLATA LAYGPAHTPP LSRIGWAMQP PPPGPIGDCL RDWEDLQODF
101 QNIQVSAAD AGSPSRVSL AOGGSGSGSPG CKPSLPAEAE GAAQELENO
151 KERQGLEFFDM EAYLPKKNGL YLSLVGNVN VTLLSKQAKF AYKDEYEKFK
201 LYTLITLILI SFTCRFLNS RVTDAAFNFL LVWYYCTLTI RESILINNGS
251 RIKGWVVFHH YVSTFLSGVM LTPDGLMYQ KFRNQLSFS MYQSFVQFLQ
301 YYQSGCLYR LRALGERHTM DLIVEGFQSW MWRVLTFLLP FLFFGHFWQL
351 FNALTFLNLA ODPCKEQOV LMGCFPELLL FLGNFTTLR VVHHKFSQR
401 HGSKKD
11AA_SEQUENCE 1.0
ID AAM84808 standard; Protein; 36 AA.
XX
AC AAM84808;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:12401.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483426/52.
 DR N-PSDB; AAK57589.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PT
 XX
 PS Claim 11; SEQ ID NO 12401; 3071pp + Sequence Listing; English.
 PS
 XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosstatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 36 AA;

AAM84808 Length: 36 May 30, 2002 09:25 Type: P Check: 1969 ..

1 RPVIMCGYCG LIFCCCFSSF WKLFRTGLX VXSRL

11AA_SEQUENCE 1.0

ID AAM93544 standard; Protein; 280 AA.

AC AAM93544;

DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 3297.

KW Human; full length cDNA; cDNA synthesis; oligo-capping.

OS Homo sapiens.

PN EP1130094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

PA (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI; 2001-524255/58.

DR N-PSDB; AAK94474.

PT 830 Primers useful for synthesizing full length cDNA clones and their

PT use in genetic manipulation -

PS Claim 8; SEQ ID NO 3297; 1380bp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been

CC isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA

CC molecules have been determined. Primers for synthesizing the full length

CC cDNA are useful for clarifying the function of the protein encoded by

CC the cDNA. The full length clones were obtained by construction of full

CC length enriched cDNA libraries that were synthesised by the oligo-capping

CC method. The primers enable the production of the full length cDNA easily

CC without any special methods. The present sequence is a polypeptide

CC encoded by a full length human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 280 AA;

AAM93544 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..

1 MQAKYSTRD MLDDGDTM SLHSQASATY RHPEPRTEH RAPSSTWRPV

51 ALTLTLCLIV LLIGLALGL LFFQYQLSN TGDPTISQME ERLGNTSQEL

101 QSLQVONIKL AGSLQHYAEK LCRELYNKAG AHRCSPCTEQ WKWHDNCYQ

151 FYKDSKSWED CKYFCLSENS TMLKINKQED LEFAASQSYS EFFYSYWTGL

201 LRPDSGKAWL WMDGTFPTSE LFHIIIDVTS PRSRDCVAAL NGMIFSKDC

251 ELKRCVCERR AGMWKPESLH VPPELTGEGD

11AA_SEQUENCE 1.0

ID AAO01590 standard; Protein; 113 AA.

XX

AC AAO01590;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 15482.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

DR N-PSDB; AA181521.

PT Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukaemia, inflammation and immune

PT disorders -

XX Claim 20; SEQ ID NO 15482; 1399pp + Sequence listing; English.

CC The invention relates to human polynucleotides (AA179941-AA193841) and

CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 113 AA;

AAO01590 Length: 113 May 30, 2002 09:25 Type: P Check: 9232 ..

1 NNFLQLKFP PKKITTDSY KAPRPFEPF PSQKGVFPPT LFFGFPFPGF

51 PPFLNPPPG FFCFWAPLKK FFFPYPGGXT WSLKGPPLR FFFFFFXDGV

101 SLCHPGWSSS AQS

11AA_SEQUENCE 1.0

ID AAB47415 standard; Protein; 585 AA.

AC AAB47415;

DT 17-OCT-2001 (first entry)

DE Plant phosphate transporter, LePT3.

KW Plant; phosphate transporter; LePT3; solanaceae; tomato;

KW phyto-remediation; pollution; phosphate-based fertilizer;

KW phosphate sink.

XX

OS Lycopersicon esculentum.
XX
FH Key Location/Qualifiers
FT Domain 18..37
FT Domain /note="Membrane spanning domain I"
FT Domain 71..92
FT Domain /note="Membrane spanning domain II"
FT Domain 100..120
FT Domain /note="Membrane spanning domain III"
FT Domain 124..144
FT Domain /note="Membrane spanning domain IV"
FT Domain 171..191
FT Domain /note="Membrane spanning domain V"
FT Domain 211..232
FT Domain /note="Membrane spanning domain VI"
FT Modified-site 243..245
FT /note="Casein kinase II phosphorylation site"
FT Domain 292..312
FT /note="Membrane spanning domain VII"
FT Domain 346..368
FT /note="Membrane spanning domain VIII"
FT Domain 372..390
FT /note="Membrane spanning domain IX"
FT Domain 399..419
FT /note="Membrane spanning domain X"
FT Domain 441..462
FT /note="Membrane spanning domain XI"
FT Domain 472..491
FT /note="Membrane spanning domain XII"
FT Modified-site 498..501
FT /note="Protein kinase C phosphorylation site"
PN WO200155299-A2.
PD 02-AUG-2001.
XX
PF 26-JAN-2001; 2001WO-IL00075.
XX
PR 27-JAN-2000; 2000IL-0134258.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Levy A;
XX
DR WPI; 2001-488780/53.
DR N-PSDB; AAH43180.
XX
PT Phosphate transporter gene encoding a phosphate transporter, Lept3,
PT useful for generating plants which can be cultivated on
PT phosphate-deficient soil, and for phyto-remediation of phosphate
PT polluted soils -
XX
PS Claim 1; Fig 1; 59pp; English.
XX
CC This sequence shows a novel plant phosphate transporter, Lept3.
CC Lept3 is useful for increasing the uptake of phosphate by a plant
CC (belonging to the family solanaceae, preferably a tomato plant).
CC Lept3 cDNA is useful for producing a plant over expressing Lept3
CC which is useful for effecting phyto-remediation of an area polluted
CC with phosphate (a portion of which is radioactive) which involves:
CC (a) providing a plant (which belongs to the family solanaceae,
CC preferably a tomato plant) over expressing Lept3 to facilitate uptake
CC and concentration of phosphate within the plant cells;
CC (b) planting the plant in the area polluted with phosphate;
CC (c) following a time period, in which at least a fraction of the
CC phosphate in the area has been accumulated in the plant, harvesting the
CC plant, thereby removing at least the fraction of the phosphate from the
CC area; and
CC (d) optionally repeating steps the second and third steps until a
CC sufficient amount of the phosphate has been removed from the area.
CC Inactivating expression of the endogenous phosphate transporter gene
CC generates plants suitable for growth under high phosphate conditions.
CC Lept3 when expressed in plants increases the ability of the plants to

CC absorb phosphate in soil, either for purposes of phyto-remediation, or
CC as a means of reducing the need of phosphate-based fertilizers.
CC Therefore Lept3 DNA is useful for generating plants which can be
CC cultivated on phosphate-deficient soil and for phyto-remediation of
CC phosphate polluted soils. Plants which accumulate high concentrations
CC of phosphate due to Lept3 serve as an excellent phosphate sink in
CC phosphate polluted areas.
XX
SQ Sequence 585 AA;
AAB47415 Length: 585 May 30, 2002 09:25 Type: P Check: 6481 ..
1 MASDNLVLN ALDTARTQWY HTAVIAGM GFTDAYDLF CITTISKLLG
51 RLYYDPPTTH APGKLPHVAN NWVIGVALVG TLSGQLVFGW LGDKLGRKKV
101 YGLTLILMWL CALCSGLSLG YSAKGVIQTL CFFRFWLFGG IGDYPLSAT
151 IMSEYANKAT RGAFTAAVFA MQGVGIIFAG LVSMIISKLF LMRYEGBEPN
201 VDEILSTEPQ ADYVWRIVLM LGALPALITV YWRMKMPETG RYTAIIEGNA
251 KQAAINMGKV LDIEIQAESD KIAQFKANE YSLLSNEFFQ RHGLHLIGTM
301 STWFLLDIAF YSQNLTKDKI FPPVGLTISNA NTISALREMF ETSRAMEVIA
351 LEGTEPGYWF TVEFIEKIGR FRIQLMGFEM MSVFMAIGV KYDYLSKEH
401 KWTFATLYGL TEEFANFGPN STTFVLPAEL FPTRVRSTGH ALSAASGRAG
451 AMISAFGIQO YTQDGNVHKX QDSYDTVCY KYGWILLPEL VTETKGRSLE
501 EITGEDGGQM RHRXRLANLS VSIKTTGWEX YKQNFIVIH HYALCFSTFN
551 YLYEXELXHY EXPQYNKLTN TNSIVXSLHR HEYLF
!!AA_SEQUENCE 1.0
ID AAM15727 standard; Protein; 27 AA.
XX
AC AAM15727;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #2161 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX

PS Claim 27; SEQ ID NO 20553; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 27 AA;

AAM15727 Length: 27 May 30, 2002 09:25 Type: P Check: 8372 ..

1 SLKKIDRPLF KDFWQFLDS LKALAVK

IIAA_SEQUENCE 1.0

ID AAM28232 standard; Protein; 27 AA.

AC AAM28232;

DT 17-OCT-2001 (first entry)

DE Peptide #2269 encoded by probe for measuring placental gene expression.

KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta -

PS Claim 27; SEQ ID NO 28501; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs:
XX see AAI31315-AAI57546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders.

AAM28232 Length: 27 May 30, 2002 09:25 Type: P Check: 8372 ..

1 SLKKIDRPLF KDFWQFLDS LKALAVK

IIAA_SEQUENCE 1.0

ID AAM39779 standard; Protein; 80 AA.

AC AAM39779;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2924.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR WPI; 2001-442253/47.

DR N-PSDB; AAI58935.

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

PS Example 4; SEQ ID NO 2924; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and

XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX immunosuppressant and cytoskeletal activity. The polynucleotides are useful

XX in gene therapy. A composition containing a polypeptide or polynucleotide

XX of the invention may be used to treat diseases of the peripheral nervous

XX system, such as peripheral nervous injuries, peripheral neuropathy and

XX localised neuropathies and central nervous system diseases, such as

XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.

AAM39779 Length: 80 May 30, 2002 09:25 Type: P Check: 8327 ..

1 MFLCAQEWLT LGINVPLLFY HFWRYFHCPA DSSELAYDPP VVMNADTISY

51 CQKEAWCKLA FYLLSFFYYL YCMITYLVSS

IIAA_SEQUENCE 1.0

ID AAM39876 standard; Protein; 531 AA.

XX AAM39876;
AC 22-OCT-2001 (first entry)
DT
XX
DE Human polypeptide SEQ ID NO 3021.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB; AAI59032.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 3021; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 531 AA;
AAM39876 Length: 531 May 30, 2002 09:25 Type: P Check: 4914 ..
1 MADVLSVLRLQ YNIOKKEIVV KGDEVIFGEF SMPKNVKTNY VVMGTGKEGQ
51 PREYTLDSI LELLNNVHLS HPVYVRRAT ENIPVRRPD RKDLLGLNG
101 EASTSASIDR SAPLEIGLOR STQVKRADE VLAELAKPRI EDEECVRLDK
151 ERLARLEGH KEGIVOTEQI RSLSEAMSVK KIAAIKAKIM AKKRSTIKTD

201 LDDDTALKQ RSVDAEVDV TRDIVSRERV WRTRTTILS TGNFESKNIF
251 AILQSVKARE EGRAPEQRP PNAAPVDPTL RTKQIPAPAY NRYDQERKRG
301 KEETEGFKID TMGTYHGMTL KSVTEGASAR KTQTPAADPV PRPVSOARPP
351 PNQKGSRTF IITIPATTS LITMLNAKDL LQDLKFVPSD EKKKGCCORE
401 NETLIQRRKD QMQPGGTAIS VTPVRYVDQ PLKLPQDMD RVAVFVQGP
451 AMQFKGWPWL LPDGSVPDIF AKIKAFHLKY DEVRLDPNVQ KMDVTYVLELS
501 YKRLHLDRPV FLRFWETLDR YMWKHKSHLR F
11AA_SEQUENCE 1.0
ID AAM40139 standard; Protein; 258 AA.
XX
AC AAM40139;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 3284.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB; AAI59295.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 5; SEQ ID NO 3284; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 258 AA;

AAM40139 Length: 258 May 30, 2002 09:25 Type: P Check: 2511 ..

1 MTVKHCALSL VGEPIWPEI NRFLKLLHQC KISSPLVTNA QFPATIRNLE
51 PVTQLVSVSD ASTKDSLKKI DRPLFKDEFWQ RFLDSLKALA VKQQRIVYRL
101 TLVKAWNYDE LQAYAQVLVL GNPDFIEVKG VTYCGESSAS SLTMAHVPMW
151 EEVVOFVHEL VDLIPEVEIA CEHEHSNCLL IAHKKFKIG EMMTWIDYNR
201 FQELIQEYED SGSKTFSAK DYMATPHWA LFGASERGF DPKDTRHQKN
251 KSKAISGC

11AA_SEQUENCE 1.0

ID AAM41565 standard; Protein; 104 AA.

AC AAM41565;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6496.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB; AAI60721.

PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 6496; 10078pp; English.

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 104 AA;

AAM41565 Length: 104 May 30, 2002 09:25 Type: P Check: 4766 ..

1 LRAVSLPLQ LVLPEYSIHS LFCIMFLCAQ EMLTLGLNVP LLFYHFWRYF
51 HCPADSSELA YDPPVVMNAD TLSYCGKEAW CKLAFYLLSF FYLLXCMIYT
101 LVSS

11AA_SEQUENCE 1.0

ID AAM41925 standard; Protein; 207 AA.

AC AAM41925;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6856.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB; AAI61081.

PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 6856; 10078pp; English.

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful

CC In gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 207 AA;
AAM41925 Length: 207 May 30, 2002 09:25 Type: P Check: 2688 ..
1 VTQLYVSVA STKDSLKKID RPLKDFWQ FLDLKLAV KQRTVRLT
51 LVKAWNDEL QAYQLVSLG NPDFIEKGV TYCGESSASS LTMAHPWHE
101 EYVQFVRELV DLIPEYEIAC EHEHSNCLLI AHRKKIGGE WWTWINYNRF
151 QELIQEYEDS GGSKTFSAKD YMARTPHWAL FGASERGFDK KDIRHQRKNK
201 SKAISGC
!!AA_SEQUENCE 1.0
ID AAU12400 standard; Protein: 280 AA.
XX
AC AAU12400;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human PRO1131 polypeptide sequence.
XX
KW Human secretory and transmembrane; PRO: mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200140466-A2.
XX
PD 07-JUN-2001.
XX
PF 01-DEC-2000; 2000WO-US32678.
XX
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 09-DEC-1999; 99US-0170262.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31243.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 10-NOV-2000; 2000WO-US30873.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI: 2001-408281/43.
DR N-PSDB; AAS21472.
XX
PT Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT lung, breast, prostate, cervical
XX
PS Claim 12; Fig 458; 813pp; English.
XX
CC AAU12172-AAU12446 represent novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIA. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.
XX
SQ Sequence 280 AA;
AAU12400 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..
1 MQAKYSSTRD MLDDDGDTTM SLHSQASATV RAPEPRRTEH RAPSSTWRPV
51 ALTLTLCLV LLIGLALGL LFFQYYQLSN TGQDTISQME ERLGNTSQEL
101 OSIQVQNIKL AGSLQHVAEK LCRELYNKAG AHRCSPTCEQ WKWHGDNQYQ
151 FYKDSKSWED CKYFCLSENS TMLKINKOED LEFAASQSYS EFFYSYWTGL
201 LRPSGKAWL WMDGTPFTSE LFHIIIDVTS PRSRDCVAIL NGMIFSKDCK
251 ELKRCVCERR AGMYKPESLH VPPELTGEGD
!!AA_SEQUENCE 1.0
ID AAG92957 standard; Protein: 592 AA.
XX
AC AAG92957;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 6711.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.

XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-0127688.
XX
XX 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
DR N-PSDB; AAH68176.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
XX
PS Claim 17; SEQ ID NO: 6711; 246bp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from Coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
XX Sequence 592 AA;
AAG92957 Length: 592 May 30, 2002 09:25 Type: P Check: 4387 ..
1 MVSLEPLASL LTTRLATLKP ALKPATHLAS LGAOYIAELV PGIRMSPNRR
51 RILPANMGAG FIGAEIAMMW ALSPSLLPKP WWVTAANLAV LQAVGHAAT
101 GIHSILPRTN RRVSRKIYNA THATGATIL TTVVGLIRH RTOIRLIGOK
151 NFGPKETIAG ISVGTLYGYA LLITGELTQH SINEVKLLIE RFLPPWISFI
201 AAVSVITLTT LTLADRVLLR RILHNSAIOA AHLNRMVFPG TEQPWEPERS
251 GSPWSEYKMG AVGSOGRAVL SGGPRKDDII TYTRLSDTET HEPRIKIFIGM
301 VPGRLSDQV DLYIHEMRRT GALRRDHIYI NNSTGTGWIT DMSAHTFEFL
351 TGGNCVTISM QYSTLPALS WYKNDGPIN AARMILIDAVL HELDQLPTGS
401 RPKFLAGES LGAYGLAEVW GDVEKLLGTA DGYLLSGAPR FSDAMNALRT
451 RRDASSSERL PVIDSGRHIR FAGEPEHLDL PATWQFPFPMI VAQHASDPYV
501 WMNAELFIRR PEMLKTPKOD HQDYFPRLRW MPEVTGWQVA LDLFTSTSV
551 GGCHGNTHHE FIDYWAALLD REYVPELRHS IAYWIRANHI KR
IIAA_SEQUENCE 1.0
ID AAG92978 standard; Protein; 507 AA.
XX
XX AAG92978;
XX
XX 26-SEP-2001 (first entry)
XX

DE C glutamicum protein fragment SEQ ID NO: 6732.
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
XX Corynebacterium glutamicum.
XX
XX EP1108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-0127688.
XX
XX 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
DR N-PSDB; AAH68197.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
XX
PS Claim 17; SEQ ID NO: 6732; 246bp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from Coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
XX Sequence 507 AA;
AAG92978 Length: 507 May 30, 2002 09:25 Type: P Check: 2649 ..
1 MGRMOSERLT PSDLVGCRYR QVQRINPPEI SPLPATMQR ARREVGLETV
51 LDRLPRQPKK RGRIPFTRAD LDNDAEIAEF ETLEAIAAGD TLITGAVFTG
101 TLEGVAMEVQ VDVLYRNPDG SYMPYVSNH RVARPDPHKT MGIATVTRLG
151 LGQPLEIKAT LRHHTIDGYR LTLALMGLEB AGAAPRESSIG AAVGQDRDWA
201 YLVDTVTRYAP AAHRALLTPA PTAAPRVKEC ATCRFPWKCE PELKAADDIS
251 LFLTGRADT YREKINTTT ALIDANLGEI SHVAAAWRAE IPYLRRRAHT
301 SAPRDEVID VDVEAYLDLG AYLMGAWDGK TYIPEVIWSD LGAAEGENE
351 ARFWMLKSR RDKARQOGOT FGVCYASNG ENHMLSTAR RFEKGKVGVP
401 DEQEIRSFIS SDQWDMFAV ARSQLVGPGG LGIKQLAPAA GFHWEEEDFA
451 GEDSLHAYLI ASTAAPEAE AARAQLLSYN GDDCRATAAV RHWLRQGART
501 APVLGNI
IIAA_SEQUENCE 1.0

ID AAG74295 standard; Protein; 66 AA.
 XX
 AC AAG74295;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen protein SEQ ID NO:5059.
 XX
 KW Human: colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma.
 XX
 OS Homo sapiens.
 PN WO200122920-A2.
 XX
 PD 05-APR-2001.
 XX
 PE 28-SEP-2000; 2000WO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 DR WPI; 2001-235357/24.
 DR N-PSDB; AAH33726.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 PS Claim 11; Page 6783-6784; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 66 AA;
 AAG74295 Length: 66 May 30, 2002 09:25 Type: P Check: 942 ..
 1 TKLHFQGGGL GNXLIVKSCN TSVQVNISGP CFPSCMHEL FFMHMGAGS
 51 WKNLPVGILG XTWACL
 11AA_SEQUENCE 1.0
 ID AAG89196 standard; Protein; 160 AA.
 XX
 AC AAG89196;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 316.
 XX
 KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
 KW GENSET.

XX
 OS Homo sapiens.
 XX
 PN WO200142451-A2.
 XX
 PD 14-JUN-2001.
 XX
 PE 07-DEC-2000; 2000WO-IB01938.
 XX
 PR 08-DEC-1999; 99US-0169629.
 PR 06-MAR-2000; 2000US-0187470.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
 XX
 DR WPI; 2001-367870/38.
 DR N-PSDB; AAH64799.
 XX
 PT Full length GENSET human nucleic acids encoding potentially secreted
 PT proteins, useful in gene therapy and vaccination against a variety of
 PT diseases, and for diagnosis of those diseases -
 XX
 PS Claim 21; Page 841; 921pp; English.
 XX
 CC The invention relates to full length GENSET human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSET gene expression. For
 CC example, they be used to treat disorders associated with decreased
 CC GENSET gene expression by rectifying mutations or deletions in a
 CC patient's genome that affect the activity of GENSET or by supplementing
 CC the patients own production of GENSET polypeptides. Conversely,
 CC antisense nucleic acid molecules may be administered to down regulate
 CC GENSET expression by binding with the cells' own genes and preventing
 CC their expression. The sense and antisense nucleic acids may also be
 CC used as DNA probes in diagnostic assays to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples, and hence to
 CC determine which patients may be in need of restorative therapy.
 CC The GENSET polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSET polypeptide expression and activity. The
 CC present sequence is a GENSET polypeptide of the invention.
 XX
 SQ Sequence 160 AA;
 AAG89196 Length: 160 May 30, 2002 09:25 Type: P Check: 4447 ..
 1 MAFTFAFCY MSLVLCAAL IFFAIWHIIA FDELRTDFKS PIDQCNPVHA
 51 RERLRNIERI CFLLRKLVLP EYSIHSIFCI MELCAQEWLT LGLNVPILLEY
 101 HFWRYFHCPA DSSELAYDPP VVMNPDTLSY CQKEAWCKLA FYLLSFFYYL
 151 YCMITYTLVSS
 11AA_SEQUENCE 1.0
 ID AAU03216 standard; Protein; 483 AA.
 XX
 AC AAU03216;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Fruit fly G protein coupled receptor, DmgPCR10.
 XX
 KW Fruit fly; G protein coupled receptor; DmgPCR10;
 KW human immunodeficiency virus; HIV; cancer; Parkinson's disease;
 KW diabetes; obesity; atherosclerosis; thrombosis; stroke; renal failure;
 KW inflammation; rheumatoid arthritis; autoimmune disorder;
 KW neurological disorder; schizophrenia; manic depression; dementia;
 KW severe mental retardation; dyskinesia; Huntington's disease;
 KW Tourette's syndrome.
 XX

OS Drosophila melanogaster.
XX
PN WO200131005-A2.
XX
PD 03-MAY-2001.
XX
PF 20-OCT-2000; 2000WO-US29002.
XX
PR 22-OCT-1999; 99US-0425676.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Lowery DE, Smith VG, Kublak TA, Larsen MJ;
XX
DR WPI; 2001-316333/33.
DR N-PSDB; AAS05895.
XX
PT New Drosophila melanogaster GPCR nucleic acids and polypeptide useful
PT for inducing an immune response, for identifying homologs and for
PT treating e.g. diabetes, obesity and manic depression -
XX
XX
PS Claim 29; Page 66; 110pp; English.
XX
CC The sequence is a fruit fly G protein coupled receptor, DmGPCR10.
CC The proteins are useful for inducing an immune response against itself in
CC a mammal. The nucleic acids are useful for identifying an animal homolog
CC of DmGPCR, by screening databases or libraries. The compounds identified
CC as binding partners or modulators of GPCR binding are useful for treating
CC diseases in animals, and for control insects that are harmful or cause
CC injury to plants or animals. Diseases treated include infections (e.g.
CC viral and human immunodeficiency virus, HIV), cancer, pain, Parkinson's
CC disease, hypotension, hypertension, diabetes, obesity, attherosclerosis,
CC thrombosis, stroke, renal failure, inflammation, rheumatoid arthritis,
CC autoimmune disorders, and psychotic and neurological disorders (anxiety,
CC schizophrenia, manic depression, delirium, dementia, severe mental
CC retardation, dyskinesias, Huntington's disease or Tourette's syndrome).
CC The nucleic acids can be used for genetic mapping, and producing
CC the GPCRs. Anti-GPCR antibodies can be used in therapy, diagnostic assays
CC and for modulating GPCR activity.
XX
SQ Sequence 483 AA;

AAU03216 Length: 483 May 30, 2002 09:25 Type: P Check: 0 ..
1 MYASLMDVGO TLAARLADSD GNGANDSGLL ATGCGLEQEQ EGLALDMGHN
51 ASADGCIVPY VPVLDRPETY IYVLYTLIF IYGVLGNGTL VILFFRHRSM
101 RNINPTYILS LALADLLVIL VCVPVATIVY TQESWPFERN MCRISEFFKD
151 ISIGVSFTL TALSGERYCA IVNPLRKLQT KPLTVFTAVM IWLAILLGM
201 PSVLFSDIKS YPVFTATGNN TIEVCSPPRD PEYAKFMVAG KALVYLLPL
251 SIIGALYIMM AKRLHMSARN MPGEQSMOS RIQARARLHV ARNVVAFVVV
301 FPIGFFPYHV FELMYHFPYT AEEDPDEFWN VLRIPLKLV RPRGLYCVSG
351 VPRQHFNRYL CCICVKKRPH LRQHSTATGM MDNTSVMSMR RSTYVGGTAG
401 NLRASLHRNS NHGVGAGGG VGGGVGSGRV GSTFHRODSMP LQHGNAHGGG
451 AGGSSSGLGA GGRTAAVSEK SPINRYESGV MRY

11AA_SEQUENCE 1.0
ID AAB93312 standard; Protein; 732 AA.
XX
AC AAB93312;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:12391.
XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 12391; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 732 AA;

AAB93312 Length: 732 May 30, 2002 09:25 Type: P Check: 4753 ..
1 MDPsADTWDL FSPLISLWIN RFIYILGFAY SISLMICVQI VIKQGNLQ
51 EKSVPKRAQD LMTNGVYSLQ EKDIFFVSGV IFFGSQTGTA KGFATVLAEA
101 VTSLDLPVAI INLKEYPD PD HLIEEYTSKN VCVFLVATYT DGLPTESAEM
151 FCKWLEEASI DFRFGKTYLK GMRYAVFGLG NSAYASHFNK VCKNVDKMLW
201 MGAHRVMSR GEGDCDVVKS KHGSTEADFR AMKTFISQL QALQGERKK
251 SCGHCCKGK CESHQGSSE REEGSHQDE LHHRDTEEEE PFESSSEEF
301 GGEDHSLNS IYDVEDLGI MDHYKKEKRE KEQGEKSGL FRNMGRENEDG
351 ERRAMITPAL REALTKQGYQ LIGSHSGVKL CRWTKSMLRG RGGCYKHTFY
XX

401 GIESHRCMET TPLSLACANKC VECWRHHTNP VGTWRWKMD QPEMILKEAI
 451 ENHONMIKQF KGVPGVKAER FEEGMTVKHC ALSLVGEPIM YPEINRFLKL
 501 LHOCKISSFL VINAQPAEI RNLEPVQLY VSDASTKDS LKKIDRPLFK
 551 DEWQRFDSL KALAVKQQT VYRLTLVKAW NVDELQAYAQ LVSLGNPDFI
 601 EYKGVTYCGE SSASSLTMAH VPMHEEVQF VRELVDLIPE YEIACEHEHS
 651 NCLLIHRRKF KIGGEWWTWI NYNRFQELIQ EYEDSGSKT FSAKDYMART
 701 PHWALFGASE RGFDPKPTRH QRKNKSKAIS GC

!!AA_SEQUENCE 1.0
 ID AAB94345 standard; Protein; 561 AA.
 AC AAB94345;
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:14855.
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 PN EP1074617-A2.
 PD 07-FEB-2001.
 PE 28-JUL-2000; 2000EP-0116126.
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 PA (HELI-) HELIX RES INST.
 PI Ola T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

Claim 8; SEQ ID 14855; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesising 5602
 full-length cDNAs defined in the specification. Where a primer set
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 to the complementary strand of a polynucleotide which comprises one of
 the 5602 nucleotide sequences defined in the specification, where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 of an oligonucleotide comprising a sequence complementary to the
 complementary strand of a polynucleotide which comprises a 5'-end
 sequence and an oligonucleotide comprising a sequence complementary to a
 polynucleotide which comprises a 3'-end sequence, where the
 oligonucleotide comprises at least 15 nucleotides and the combination of
 the 5'-end sequence/3'-end sequence is selected from those defined in
 the specification. The primer sets can be used in antisense therapy and
 in gene therapy. The primers are useful for synthesising polynucleotides,
 particularly full-length cDNAs. The primers are also useful for the
 detection and/or diagnosis of the abnormality of the proteins encoded by
 the full-length cDNAs. The primers allow obtaining of the full-length
 cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 561 AA;

AAB94345 Length: 561 May 30, 2002 09:25 Type: P Check: 6460 ..

1 MRAVFGIGN SAVASHFNKV GKNVDKWLWM LGAHRVMSRG EGDGDVYKSK
 51 HGSIEADFERA WKTFISOLQ ALQGERKKS CGGCKKGGC ESHQHGSEER
 101 EEGSHQDEL HHRDTEEEP FESSSEEEFG GEDHQSLSI VDVEDLGKIM
 151 DHVAKKEKREK EQQEEKSGLF RNMGRENDEG RRAMITPALR EALTQGYQL
 201 IGSISGVKLC MWTKSMFRGR GGCYKHTFYG IESHRCMETT PSLAYANKCV
 251 FCWRHHTNPV GTEWRWKMDQ PEMILKEAIE NHONMIKQF GVPGVKAERF
 301 EEGMTVKHCA LSLVGEPIFY PEINRFLKL HOCKISSFLV TNAQPAEIR
 351 NLEPVQLYV SVDASTKDSL KRIDRPLFKD FQWRFDSLK ALAVKQQRIV
 401 YRLTLVKAWN VDELQAYAO VSLGNPDFIE VKGVTYCGES SASSLTMAHV
 451 PMHEEVQFV RELVDLIPEY EIACEHEHSN CLLIAHRRKF IGGEWWTWIN
 501 YNRFQELIQE YEDSGSGSKTF SAKDYMARTP HWALFGASER GFDPKPTRHQ
 551 RKNKSKAISG C

!!AA_SEQUENCE 1.0
 ID AAB60464 standard; Protein; 160 AA.
 AC AAB60464;
 DT 24-APR-2001 (first entry)
 DE Human cell cycle and proliferation protein CCYPR-12, SEQ ID NO:12.
 KW Cell cycle and proliferation protein; CCYPR; human; agonist;
 KW antagonist; gene therapy; detection; gene therapy;
 KW transgenic animal disease model; immune disorder;
 KW developmental disorder; cell signalling disorder;
 KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
 KW arteriosclerosis; asthma; allergy; diabetes mellitus;
 KW menstrual cycle disorder; bacterial infection.
 OS Homo sapiens.
 PN WO200107471-A2.
 PD 01-FEB-2001.
 PE 21-JUL-2000; 2000WO-US19948.
 PR 21-JUL-1999; 99US-0145075.
 PR 08-SEP-1999; 99US-0153129.
 PR 10-NOV-1999; 99US-0164647.
 PA (INCY-) INCYTE GENOMICS INC.
 PI Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
 PI Azimzal Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
 DR WPI; 2001-112727/12.
 DR N-PSDB; AAF59601.
 PT Human cell cycle and proliferation proteins and polynucleotides are
 PT used to treat, diagnose and prevent immune, developmental and cell
 PT signaling disorders and cell proliferative disorders including cancer -

Claim 1; Page 123; 205pp; English.

XX Sequences AAB60453-AAB60506 represent 54 human cell cycle and
CC proliferation proteins (CCYPR), which are encoded by AAF59590-AAF59643.
CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC associated with decreased expression of functional CCYPR, while CCYPR
CC antagonists are used to treat diseases or conditions associated with
CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC that specifically bind to CCYPR, and in drug screening methods to
CC identify compounds that modulate the activity of CCYPR. CCYPR
CC nucleotides can be used to generate transgenic animal models of human
CC disease, and can be used in gene therapy in target cells with genetic
CC abnormalities with respect to the expression of CCYPR for the
CC treatment or prevention of a disorder associated with CCYPR.
CC Diseases which can be diagnosed, treated and prevented using CCYPR
CC proteins, nucleic acids, agonists or antagonists include immune,
CC developmental and cell signalling disorders, and cell proliferative
CC disorders including cancer. Specific examples of these disorders
CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC diabetes mellitus, disorders of the menstrual cycle and infections
CC caused by bacteria.

XX SQ Sequence 160 AA;

AAB60464 Length: 160 May 30, 2002 09:25 Type: P Check: 4282 ..

1 MATFAAFCY MSLVLCAL IFFAIWHIIA FDELRTDFKS PIDQCNPVHA
51 RERLRNIERI CFLRLKLVLP EYSIHSLECI MFLCAQEWLT IGLNVPLLFY
101 HFWRYFHCPA DSSELAYDPP VVMNADTLST CQKEAWCKLA FYLLSFFYYL
151 YCMITYTLVSS

IIAA_SEQUENCE 1.0

ID AAB60465 standard; Protein; 531 AA.

XX AC AAB60465;

XX DT 24-APR-2001 (first entry)

XX DE Human cell cycle and proliferation protein CCYPR-13, SEQ ID NO:13.

XX KW Cell cycle and proliferation protein; CCYPR; human; agonist;
KW antagonist; gene therapy; detection; gene therapy;
KW transgenic animal disease model; immune disorder;
KW developmental disorder; cell signalling disorder;
KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
KW arteriosclerosis; asthma; allergy; diabetes mellitus;
KW menstrual cycle disorder; bacterial infection.

XX OS Homo sapiens.

XX PN WO200107471-A2.

XX PD 01-FEB-2001.

XX PF 21-JUL-2000; 2000WO-US19948.

XX PR 21-JUL-1999; 99US-0145075.

XX PR 08-SEP-1999; 99US-0153129.

XX PR 10-NOV-1999; 99US-0164647.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;

XX DR WPI; 2001-112727/12.

XX DR N-PSDB; AAF59602.

PT Human cell cycle and proliferation proteins and polynucleotides are
PT used to treat, diagnose and prevent immune, developmental and cell
PT signaling disorders and cell proliferative disorders including cancer -
XX
XX PS Claim 1; Page 123-124; 205pp; English.

XX Sequences AAB60453-AAB60506 represent 54 human cell cycle and
CC proliferation proteins (CCYPR), which are encoded by AAF59590-AAF59643.
CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC associated with decreased expression of functional CCYPR, while CCYPR
CC antagonists are used to treat diseases or conditions associated with
CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC that specifically bind to CCYPR, and in drug screening methods to
CC identify compounds that modulate the activity of CCYPR. CCYPR
CC nucleotides can be used to generate transgenic animal models of human
CC disease, and can be used in gene therapy in target cells with genetic
CC abnormalities with respect to the expression of CCYPR for the
CC treatment or prevention of a disorder associated with CCYPR.
CC Diseases which can be diagnosed, treated and prevented using CCYPR
CC proteins, nucleic acids, agonists or antagonists include immune,
CC developmental and cell signalling disorders, and cell proliferative
CC disorders including cancer. Specific examples of these disorders
CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC diabetes mellitus, disorders of the menstrual cycle and infections
CC caused by bacteria.

XX SQ Sequence 531 AA;

AAB60465 Length: 531 May 30, 2002 09:25 Type: P Check: 4914 ..

1 MADVLSVLRLQ YNIQKKEIVV KGDEVIFGEE SMPKNVKTNY VVMGTGKEGO
51 PREYTTLDSI LFLNNVHLS HPVYVRRAT ENIPVRRPD RKDLLGYLNG
101 EASTSASIDR SAPLEIGLQR STQVKRADE VLAEAKKPRI EDEECVRLDK
151 ERLARLEGH KEGIVQTEQI RSLSEAMSVE KIAAIKAKIM AKKRSTIKTD
201 LDDDTLALKQ RSFVDAEVDV TRDIVERSERV WRTFTTILQS TGNFESKNIF
251 AILQSVKARE EGRAPEGRPA PNAAPVDPTL RTKQPIPAAY NRYDERFKG
301 KEETEGFKID TMGTYHGWTL KSVTEGASAR KTQTPAAQPV PRPVSAQRP
351 PNQKGSRTP IILIPAAFTS LITMLNAKDL LQDLKRVPSD EKKKQCGQRE
401 NETLIQRRKD QMOPGTAIS VTPVRYVDQ PLKLMQDWD RVVAVFVQGP
451 AMQFKGWPWL LPDGSFYDIF AKIKAFHLKY DEVRLDPNVQ KMDVTYVLELS
501 YHKRHLDREV FLRFWETIDR YMVKHKSHLR F

IIAA_SEQUENCE 1.0

ID AAB65251 standard; Protein; 280 AA.

XX AC AAB65251;

XX DT 02-APR-2001 (first entry)

XX DE Human PRO1131 (UNQ569) protein sequence SEQ ID NO:319.

XX KW Human; secreted and transmembrane protein; PRO; cytosolic;

XX KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;

XX OS Homo sapiens.

XX PN WO200073454-A1.

XX PD 07-DEC-2000.

XX 30-MAR-2000; 2000WO-US08439.
PF
XX 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 07-JUL-1999; 99US-0143048.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 17-AUG-1999; 99US-0149396.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 08-OCT-1999; 99US-0158663.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gertitsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Kijavir LJ, Napier MA, Pan J, Paoni NF;
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
PI Zhang Z;
XX
XX WPI; 2001-032160/04.
DR N-PSDB; AAF44218.
XX
PT PRO polynucleotides used to produce polypeptides used to target
PT bioactive molecules such as toxins, radiolabels or antibodies, to
PT specific cells, to cause targeted cell death -
XX
XX Claim 12; Fig 230; 935pp; English.
PS
XX The present invention describes human secreted and transmembrane PRO
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
CC can be used for targeted delivery of bioactive molecules, such as
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
CC sequences, and their fragments, can be used as hybridisation probes, in
CC chromosomal and gene mapping, and in the generation of anti-sense RNA
CC and DNA. They may also be used to produce transgenic animals which are
CC used to develop and screen therapeutically useful reagents. The PRO
CC nucleotide and protein sequence can be used for tissue typing and in
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX
XX Sequence 280 AA;
SQ
AAB65251 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..

251 ELKRCVCERR AGWVKPESLH VPETLGECD
IIA_SEQUENCE 1.0
ID AAB50959 standard; Protein; 280 AA.
XX
AC AAB50959;
XX
DT 21-MAR-2001 (first entry)
XX
DE Human PRO1131 protein.
XX
KW Human; PRO; cytosolic; nootropic; neuroprotective; respiratory general;
KW antiinflammatory; antiangiogenic; immunosuppressive; immunostimulant;
KW PRO agonist; cancer; inflammatory disorder; immunological disorder.
XX
OS Homo sapiens.
XX
PN WO200073348-A2.
XX
PD 07-DEC-2000.
XX
PF 30-MAY-2000; 2000WO-US14941.
XX
PR 02-JUN-1999; 99WO-US12252.
PR 22-JUN-1999; 99US-0140650.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30999.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-0187202.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
XX
PA (GETH) GENENTECH INC.
XX
XX Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC;
PI Shelton DL, Smith V, Watanabe CK, Wood WI;
XX
DR WPI; 2001-016509/02.
DR N-PSDB; AAC91561.
XX
XX Claim 31; Fig 18; 188pp; English.
PS
XX The present sequence is one of twenty eight novel PRO polypeptides. The
CC PRO polypeptides and their agonists, including antibodies, peptides, and
CC small molecule agonists, may be used to treat various tumours, e.g.,
CC cancers such as breast cancer, ovarian cancer, renal cancer, colorectal
CC cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer,
CC central nervous system cancer, melanoma or leukaemia. They are also
CC useful for treating other disorders such as neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal and
CC blastocoelic disorders, and inflammatory, angiogenic and immunological
CC disorders.
XX
XX Sequence 280 AA;
SQ

AAB50959 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..
1 MQAKYSSTRD MDDDDGDTM SLHSQASATP RHEPRRTTEH RAPSSTWRPV
51 ALTLTLTCLV LLIGLALGL LFFQYYQLSN TGQDTISQME ERLGNTSQEL
101 OSLOVONIKL AGSLQHAVER LCRELYNKAG AHRCSPTCEQ WKWHGDNQCYQ
151 FYKDSKSWED CKYFCLSENS TMLKINKQED LEFAASQSYS EFFYSWTGL
201 LRPDGKAWL WMDGTPFTSE LFHIIIDVTS PRSRDCVAIL NGMIFSKDCK
251 ELKRCVCERR AGWVKPESLH VPPETLGEGD
11AA_SEQUENCE 1.0
ID AAU19776 standard; Protein: 394 AA.
XX
AC AAU19776;
XX
DT 06-DEC-2001 (first entry)
XX
DE Human novel extracellular matrix protein, Seq ID No 426.
XX
KW Human; secreted extracellular matrix protein; immunomodulatory;
KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
KW Sezary syndrome; Gaucher's disease; neurological diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW cardiac arrest; tachycardia; angina; infection; corneal infections;
KW wound healing; immunogen; gene therapy; antisense; food additive.
XX
OS Homo sapiens.
XX
PN WO200155368-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01348.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-02411221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465572/50.
DR N-PSDB; AAS31347.
XX
PT Nucleic acid molecules encoding human secreted extracellular matrix
PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers -
XX
PS Claim 11; SEQ ID No 426; 577pp; English.
XX
CC The invention relates to isolated nucleic acid molecules encoding
CC novel human secreted extracellular matrix proteins (SPs). The
CC polynucleotides and proteins are used to prevent, treat a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. For example, disorders associated with decreased
CC expression of SPs. The SP polynucleotide or a vector expressing them may
CC be administered to treat diseases by gene therapy. Antisense molecules
CC may be administered to down regulate expression of SPs by binding with
CC the cells own genes and preventing their expression. The polynucleotides
CC may also be used as DNA probes in diagnostic assays. The SPs may also be
CC used as antigens to produce antibodies and to identify modulators
CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
CC antagonists may also be used to down regulate expression and activity of
CC SP and as diagnostic agents for detecting the presence of SPs in samples.
CC The disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac
CC arrest, tachycardia and angina), infections caused by bacteria, viruses
CC and fungi and ocular disorders (e.g. corneal infections). Other uses
CC include wound healing, maintenance of organs before transplantation,
CC support of cell culture of primary tissues, modulation of for example
CC differentiation of embryonic stem cells, mammalian characteristics
CC (e.g. height and weight), catabolism, anabolism, energy storage,

CC mental state, biorhythms, cardiac rhythms, reproductive potential,
CC hormonal levels appetite, memory and stress. The SPs can also be used as
CC an food additive. Full details of uses diseases are disclosed in the
CC specification. The present sequence represents a novel human secreted
CC extracellular protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 394 AA;
AAU19776 length: 394 May 30, 2002 09:25 Type: P Check: 52 ..
1 DRRRPASGGP GLLRPPPLA TGAXAMQPPP PGPLGDCLRD WEDLQDFQN
51 IQETHRLYRL KIELTKLQN NCTSSITRQK KRLQELALAL KKCKPSLPAP
101 AEGAAQELN QMKERQGLFF DMEAYLPKKN GLYLSVLGN VNVTLISKQA
151 KPAVKDEYEK FKLYLTILLI LISFTCRFL NSRVTDAPFN ELLWYYCTFL
201 TIRESLINN GSRKIGWVWF HHVSTFLSG VMLTPPDGLM YQFRNQFLS
251 FSMYQSFVQF LQYYQSGCL YRLRALGERH TMDLTVEGFQ SWMWRLTFL
301 LPFLFFGFHW QLFNALTFFN LAQDPQCKEW QGCAPOVSQS AAREQEGRL
351 GLPLPAQRGF CPVCCGRGWE APLEACIRG SLFLSGFYG RCGP
11AA_SEQUENCE 1.0
ID ABB48814 standard; Protein; 510 AA.
XX
AC ABB48814;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #1518.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihl H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjarl H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kunh M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, de Pablos B, Wehlend J, Kaerst U, Entlian K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides -
XX
PS Claim 6; SEQ ID No 1519; 192pp; French.
XX
CC The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of

CC It are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 510 AA;
ABB48814 Length: 510 May 30, 2002 09:25 Type: P Check: 2157 ..
1 MSKSPVAIIII LDGFGKRAET VGNAVAQANK PNFDRYWADF PHGELKAAGL
51 DVGLEPGQMG NSEVGHNTNIG AGRIVYQSLT RIDKAIEEGE FOENKALNNA
101 FTHTKENNSD LHLFGLSDG GVHSHINHLV ALLETAKDKG VKNVYIHAFL
151 DGRDVAPOSS LEYLETLQKA ISDLNYGAIA TVSGRFYAMD RDKRWERVEK
201 AYKAIVSAEG EKFEDEPIELV KASYANDKND EFVVPALITK DGKPVATVKD
251 NDAVIEFNER PDRAIQLSNA FTDKEWDHFD RGADHPKNIK FVTMTLYNPS
301 IDAEVAFEPPI EMKNVIGEVN SNEGLSQLRI AETEKYPHVT FFMNGRNEE
351 FPGENRILIN SPKVETTYDLQ PEMSAVEYTD ALVEDIKNDK HDAIILNFAN
401 PDMVGHSGML EPTIKAIEAV DENLGRVVDL ILEKGSAIL FADHGNSETM
451 STPECKPHTA HTTVVPVPIV TKKGVTLREG GRLADVAPTM LDLLGVKKPA
501 EMTGESLIQK
11AA_SEQUENCE 1.0
ID ABB49398 standard; Protein; 289 AA.
XX
AC ABB49398;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #2102.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst P, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Doman E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entlan K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides -
XX
PS Claim 6; SEQ ID No 2103; 192pp; French.
XX
CC The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms. The present sequence is a protein
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 289 AA;

ABB49398 Length: 289 May 30, 2002 09:25 Type: P Check: 1607 ..
1 VMKKVVKYK QNCIVQVGA VSARVGRNDV SGNAQLAYY MLFSIFPMLL
51 IATLLAYLH IDKDSVFNMI KEFAPDQIMD FLEENLNML TOKNGGLSTI
101 GIATLMSAS NGMNAVMKSL NKAYGVTKNR NYVQRLSM FETLAMLATV
151 GATLLLVFG QOIGMFLINH LNFSDFLSF WNNLRWTVTL IVIFVVFTEL
201 YWVAPNRRST LISVLPALF STIGWTVASV GFAYVYNNFG NYSATYGSIG
251 VITILLMWFY LTGIILMIGG ELNATLAIK KKKELGEIN

This Page Blank (uspto)

11AA_SEQUENCE 1.0
P1;G71532 - probable threonine/tyrosine-specific protein kinase (EC 2.7.1.-) -
Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: G71532
R:Stephens, R.S.; Kalman, S.; Jammal, C.J.; Fan, J.; Marathe, R.; Aravind, L.;
Mitchell, W.P.; Olinger, L.; Tatusov, R.L.; Zhao, Q.; Koonin, E.V.; Davis, R.W.
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans:
Chlamydia trachomatis.
A:Reference number: A71570; MUID:99000809
A:Accession: G71532
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-934 <ARN>
A:Cross-references: GB:AE001302; GB:AE001273; NID:93328708; PIDN:AAC67894.1;
PID:93328716
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: pknD
C:Superfamily: Chlamydia trachomatis probable threonine/tyrosine-specific
protein kinase; protein kinase homology
C:Keywords: phosphotransferase
F;2-293/Domain: protein kinase homology <KIN>

G71532 Length: 934 May 30, 2002 09:25 Type: P Check: 4900 ..

```
1  MQRVELRLI GKGGMGEVYL AHDKACSRV ALKRIREDL GNALLRQFL
51  REAKIADLI HPGIVPVYSI CSDGEAVYTT MPYIEGSLK SLKSWQKE
101 VLSKELEKT SVKSFLLPFD KICATVEYIH SKGYLHDLK PDNILLGLFG
151 EVVIIDMGAA IFKHAKEKL EQDDEAVSF DERNICYSSM TTPGKIYGT
201 DYMAPESLIG VEASEKTDIY ALGLILYQL TLAFFYRRKK GRKLSRDYV
251 LPIEMSPYR EIPPSLSOIA MKAIINPAD RFSIQELRQ ALQPYLGDP
301 EMTVKATLMA KEKSCWKYYD PILSRYPV LASSPAQWYN FMLSEVEISA
351 STRVEYVTK SAVHEGMGL FLPSKEAERG EFYCGYGLWF SVQNHETVS
401 LIKNGIETOK KSQEMISQOS RFALIEKSD NRIAVFEQA LFIHLIDYLP
451 SLGNRLGVII QDLQGSNTA ISESIGALRV SCLAVPADFL SEKLYDQAI
501 FYRKIRDSFP GRKESYEAF RLGVTLTQI EEQGGDLTQA LSSFDYHGG
551 AGAPLEYLCK ALVYQRNSF VEEIRCLLFA LKRYSQHPEI PRLEDHLCFR
601 LYDSLHKHRS EALVFMILL WIAPEKISVR EEEFRLRIY HKQATLFCQ
651 VDKAPLQFRS SKMELFLSFW TGFSLFLPEL FRRAGGLDY QALADIFYVA
701 GVSQNRFAFM QESTALANVS DEITFPESIH NQKVAELMEF VKGVEALRNK
751 DYQAKAKLLM KTFPTQLYA LDMFHQAFI DEIESFIDL LQAIYDPASE
801 EERDHLIYI IQTHLMNRDL ERAYKLLNDR FPLDEELAEY SEAFILWCY
851 LALTGDRVYV KAHFSRCRYK YKSALIGKC VDGDIIFYLD NLVWWEKMT
901 LFQSYFLRC LNESPRRYEK YRQAYLSMEN NFFD
```

11AA_SEQUENCE 1.0
P1;W5WLR1 - E5 protein - rhesus papillomavirus (type 1)
C:Species: rhesus papillomavirus
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 27-Jan-1995
C:Accession: F38503
R:Ostrow, R.S.; Labresh, K.V.; Faras, A.J.
Virology 181, 424-429, 1991

A:Title: Characterization of the complete RbPV 1 genomic sequence and an
integration locus from a metastatic tumor.
A:Reference number: A38503; MUID:91135018
A:Accession: F38503
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-157 <OST>
A:Cross-references: EMBL:M37717
C:Superfamily: rhesus papillomavirus E5 protein
C:Keywords: early protein

W5WLR1 Length: 157 May 30, 2002 09:25 Type: P Check: 8739 ..

```
1  IVVICGTQWS HKPVHTLNS IQVLCKANCC CYACKPPPC CFWLCECCF
51  CLALCFVHLL SRCFCVFPVC LSVAAVAVYL GVHSEPVCSF WSVFVLEFPN
101 VAFDTPACPQ CGLQQNDVNT AHRVITISYF AIVAVNIYFV LALLVGAAFK
151 ATSRART
```

11AA_SEQUENCE 1.0
P1;C69762 - di-tripeptide ABC transporter (membrane pr) homolog yclF - Bacillus
subtilis
C:Species: Bacillus subtilis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: C69762
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo,
V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boriss, R.;
Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani,
J.J.; Conerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.;
Duesterhoeft, A.; Ehlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.;
Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi,
A.; Galleron, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi,
G.; Guiseppi, G.; Guy, B.J.; Haga, K.; Hahsch, J.; Harwood, C.R.; Henaute, A.;
Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Itaya, M.; Jones, L.;
Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.;
Kobayashi, Y.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.;
Lapidos, A.; Lardinois, S.
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda,
S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.;
Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogilwara, A.; Oudega,
B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Porwolik, S.; Prescott,
A.M.; Prescan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds,
S.; Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato,
T.; Scanlon, E.
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska,
A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Tacconi, E.;
Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka,
T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenbol, M.;
Vannier, F.; Vassarotti, A.; Viari, A.; Wambutt, R.; Wedler, E.; Wedler, H.;
Weltzenegger, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis.
A:Reference number: A69580; MUID:98044033
A:Accession: C69762
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-492 <KUN>
A:Cross-references: GB:Z99106; GB:AL009126; NID:92632653; PIDN:CAB12175.1;
PID:92632668
A:Experimental source: strain 168
C:Genetics:
A:Gene: yclF
C:Superfamily: peptide transporter protein

C69762 Length: 492 May 30, 2002 09:25 Type: P Check: 9486 ..

1 MASIDNESII KSPQKGFFG HPRGLFTLEF TEFWERSY GMRAILEYLL
51 YTETVNGGLG FDKGTAVAIM SIYGLVYMS TIIGWLADR VEGTANTVEY
101 GGIFIMEGHI ALAYPCSSIA FYISWLIIV GTGLKPNVS SVGDLYTKE
151 DPRDSGFSI FYMGINLGL LAPLIVGILG QKYNHILGFG AAAYGMLLGL
201 IVFPLTRKN LGLAGSNVPN PLSKSAIGT GIGVIVAIA VIISVOTGVL
251 TIKRFIDIVS ILGILIPVIY FIIMFTSKKA DKTEKSRLAA YVPLFGAVM
301 FWAIOEGAT ILAVYADERI RLSLGFELQ SSWFQSLNPL FVIFADIFA
351 WLMMKLGKRO PSTPVKFSIG IILAGLSRII MVFPAMQKE ALVSPMLVL
401 SELLVLGEL CLSPVGLSVT TKLAPAFSA QTMSEMLTN AAAQAINAOV
451 AGLFDKIPET MYFGTIGLIS IVLGGILLLL SPVIRKAMKG VL

11AA_SEQUENCE 1.0

P1;C69291 - pheromone shutdown protein (trab) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: C69291
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.; Peterson, J.D.; Richardson, D.L.; Kerlavage, A.R.; Graham, D.E.; Kyrpides, N.C.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Dougherty, B.A.; McKenry, K.; Adams, M.D.; Loftus, B.; Peterson, S.; Reich, C.I.; McNeil, L.K.; Badger, J.H.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.; D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.; Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Moese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Archaeoglobus fulgidus.
A;Reference number: A69250; MUID:98049343
A;Accession: C69291
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-396 <KLE>
A;Cross-references: GB:AE001082; GB:AE000782; NID:g2689405; PIDN:AAB90915.1; PID:g2650316; TIGR:AF0331
C;Superfamily: pheromone shutdown protein

C69291 length: 396 May 30, 2002 09:25 Type: P Check: 4349 ..

1 MEEKRLVVG TAHVSKRSVE EVAEYTEREK PDAVAVELCP RRYHALVHGQ
51 REEISVADVI RKGNVMLLF QLILAYFQRK VGEETGVKPG SEMLAIEKA
101 REAGADVLLI DRDIGLFTFTR FWQKLFVEK IKLIFHLVRS TFSGDEIEVD
151 EMLEEDVLDL LVKEFRKISP NAAKVLIDER DVMANLNLN ALSRYNRIVA
201 VVGAGHRKGI EEALTKKEN PVNLRELEEV KKGRSYFKIF MGAFYALIVG
251 TFIILATSLN TEVLVQAFLY WFLINGVLSA VGASLARGHP LSIYTAFLCA
301 WMTSLNPLVA AGWVSGLYEA WIRKPYEDEF SKIVEASLR EMMQNKFERV
351 LLVAALNVG SMIGTIYGW YIISTFGVDV AKVGERVVE VIGGLL

11AA_SEQUENCE 1.0

P1;G86502 - S/T protein kinase [imported] - Chlamydia pneumoniae (strain J138)
C;Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: G86502
R;Shirai, M.; Hiraoka, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ishii, K.; Hattori, M.; Kuhara, S.; Nakazawa, T.

Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349
A;Accession: G86502
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-918 <STO>
A;Cross-references: GB:BA000008; NID:g8978468; PIDN:BAA98305.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: CPj0095
C;Superfamily: Chlamydia trachomatis probable threonine/tyrosine-specific protein kinase; protein kinase homology

G86502 length: 918 May 30, 2002 09:25 Type: P Check: 36 ..

1 MGEVYLAYDP VCSRKVALKK IREDLAENPL LKRRFLREAR IAADLIHPGV
51 VPVYTIYSEK DPVYTYMPYI EGYTLKTLIK SVWQKESISK ELAEKTSVGA
101 FLSEFHKICC TIEYVHSRGI LHRDLKPDNI LGLFSEAVI LDWGAANVACG
151 EEDLDLDIV SKEEVLSSRM TIPGRIVGTP DYMAPERLLG HPASKSTDIY
201 ALGVVLYOML TLFPPYRRK GKKIYLDGOR IPSPOEVAVY REIPPELSAV
251 VMRLAVDPQ ERYSSVTELK EDIESHLKGS PKWTLTVALP PKSSSWKLN
301 EPIILSKYRP MLEVSPASWY SLAISNIESF SEMRLEYTLS KGLNEGFGI
351 LLPTSENALG GDFYQYGFV LHIKERTLSV SLVKNSTLEIQ RSCQDLES DK
401 ETEFLALEQH NLSLSLFVDG TTWLHMYNL PSRSGRAVLI VRDMEDILED
451 IGIFESSGSL RVSCLAVPDA FLAEKLYDRA LVLYRRIAES FPGKRGYEA
501 RFRAGITVLE KASTDNEGE FALAIEERSK LHDGVAAPLE YLGKALVYQR
551 LOEYNEEIKS LLLALKRYSQ HPEIFRLKDH VVYRLHSEFY KRDRLALVEM
601 ILVLEIAPQA ITPGQEEKIL VWLKDKSRAT LFCLLDPTVL ELRSSKMELEF
651 LSYWSGFIPH LNSLFHRAMD QSDVRALIEI FYVACDLHKW QFLSSCIDIF
701 KESLEDQKAT EEIVEFSFED LGAFLEAIQS IFNKEDAEMI FVSNQQLSPI
751 LLVYIFDLFA NRALLESGE AIFQALDLIR SKVPENFYHD YLRNHEIRAH
801 IWCNRKALS TIFENYTERQ LKDEQHLELV LYGCYALAIQ GAENAKQHED
851 VCREDRIEFA SLARNYNRL GLPKDALSYQ ERLLLRQKF LYFHCLGNHD
901 ERDLQQTMYH LLTEEFOL

11AA_SEQUENCE 1.0

P1;C72120 - s/t protein kinase - Chlamydia pneumoniae (strain CWL029)
C;Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C;Accession: C72120
R;Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Davis, R.W.; Stephens, R.S.
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606
A;Accession: C72120
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-918 <ARN>
A;Cross-references: GB:AE001595; GB:AE001363; NID:g4376353; PIDN:AAD18248.1; PID:g4376355
A;Experimental source: strain CWL029
C;Genetics:
A;Gene: Cpn0095

C:Superfamily: Chlamydia trachomatis probable threonine/tyrosine-specific protein kinase; protein kinase homology

C72120 Length: 918 May 30, 2002 09:25 Type: P Check: 36 ..

1 MGEVYLAYDP VCSRKVALKK IREDLAENPL LKRRFLREAR IAADLIHPGV
51 VPVYTIYSEK DPVYITMPYI EGYTLKTLK SWQKESLSK ELAEKTSVGA
101 FLSIFHKICC TIEYVHSRGI LHRDLKPDNI LGLFSEAVI LDWGAAYACG
151 EEEEDLDIDV SKEEVLSSRM TIPGRIVGTP DYMAPERLLG HPASKSTDIY
201 ALGVVLYQML TISFPYRRKK GKRIVLGQR IPSPOEVAPY REIPPELSAV
251 VMRLAVDPQ ERYSSVTELK EDIESHLKGS PKWTLTALP PKSSSWKLN
301 EPIILSKYFP MLEVSPASWY SLAISNIESF SEMRLEYTIS KKGINEGFCI
351 LLPTSENALG GDFYQGYGFW LHIKERTLSV SLVKNLSLEIQ RCSQDLES DK
401 EETFLALEQH NHSISLFVDG TFWLIHMNYL PSRSGRAVAI VRDMEDILED
451 IGIFESSGSL RVSGCLAVPDA FLAEKLYDRA LVLYRRIAES FPGKEGYEA
501 RFRAGITVLE KASTDNNEOE FALAIEEFSK LHDGVAAPLE YLGKALVYQR
551 LOEYNEEIKS LLLALKRYSQ HPEIFRLKDH VYRLHESFY KRDRALVEM
601 ILVLEIAPQA ITPGQEEKIL VWLKDKSRAT LFCLLDPTVL ELRSSKMELF
651 LSWYSGFIPH LNSLFHRAWD QSDVRALEI EYVACDLHKW QELSSCIDIF
701 KESLEDQKAT EEIVEFSFED LGAFLEAIQS IFNKEDAERI FVSNQDLSPI
751 LLYVIFDLFA NRALLESQGE AIFQALDLIR SKVPENFYHD YLRNHEIRAH
801 LMCNNEKALS TIFENYTEKO LKDEQHLEFV LYGCYLALIQ GAERAKQHFD
851 VCREDRIFPA SLIARNYNRL GLPKDALSYQ ERRLLLRQKF LYFHCIGNHD
901 ERDLQOTMYH LITREFOQL

!!AA_SEQUENCE 1.0

P1:AF1811 - dtdp-glucose 4-6-dehydratase [imported] - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.

A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AF1811

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Itiguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120.

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AF1811

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-356 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAB77562.1; PID:g17135016; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: rfbB

C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

AF1811 Length: 356 May 30, 2002 09:25 Type: P Check: 1483 ..

1 MIONDETLT TSVAAMPILI TGAGAGIGSN FVHHWEKYP GDRMIVLDAL

51 TYAGNRONLA DIEKANLRF VKGIDIGDRAL IDQLLEEKI QAIHFAES
101 HYDRSIYAPD AFIQTNVGT FTLLAEFRHH WTKQKRPAN RELHVSTDEV
151 YGSLIEDDPA FTEPTPYAPN SPYSASKAGS DHLARAYYHT YGLPTLITNC
201 SNNYGPYHP EKLIPPLICLN ILLGRPLPIY GDGINIRDWL YVEDHCRALD
251 IVIHQKPE TYNIGNNEI KNIDLQMIC ELMDELAPDL PVPASKLIT
301 FVKDRPGHDR RYAINATKIK TELGWEPOQT ISTGLRHTIQ WYLTHRHWE
351 ALLPKE

!!AA_SEQUENCE 1.0

P1:AH1381 - phosphoglycerate mutase homolog pgm [imported] - Listeria monocytogenes (strain EGD-e)

C:Species: Listeria monocytogenes.

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AH1381

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fstli, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordstiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AH1381

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-510 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAD00534.1; PID:g16411944; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: pgm

C:Superfamily: phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent

AH1381 Length: 510 May 30, 2002 09:25 Type: P Check: 2157 ..

1 MSKSPVAIII LDGFGKRAET VGNVAQAANK PNEDRYWADF PHGELKAAGL
51 DVGLPEGQMG NSEVGHNTIG AGRIYQSLT RIDKAIEEGE FOENKALNNA
101 FTHTKENNSD LHLFGILSDG GYHSHINLIV ALLETAKDKG YKNYIHAFL
151 DGRDVAPQSS LEYLETLQKA ISDLNYGAIA TVSGRFYAMD RDKRWERVEK
201 AYKAIVSAEG EKFEDEIELV KASYANDKND EFVPAIITK DGRPVATVKD
251 NDAVIFENFR PDRAIQLSNA FTDKEMDHF D RGADHPKNIK FYTMTLYNPS
301 IDAEVAFEPI EMKNVIGEVL SNEGLSQLRI AETEKYPHYT FFMNGRNEE
351 FPGENRILIN SPKVETVDLQ PEMSAVEVTD ALVEDIKNDK HDAIILNPN
401 PDMVGHSGML EPTIKAIEAV DENLGRVVDL ILEKGSALII FADHGNSETM
451 STPEGKPHTA HTTVVPVIV TKKGVTLLREG GRLADVAPTM LDLIGVKKPA
501 EMTGESLIQK

!!AA_SEQUENCE 1.0

P1:AI1750 - phosphoglycerate mutase homolog pgm [imported] - Listeria innocua (strain C1p11262)

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C;Accession: A11750
 R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Rimmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.
 A;Title: Comparative genomics of *Listeria* species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: A11750
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-510 <GLA>
 A;Cross-references: GB:AL592022; PIDN:CAC97777.1; PID:g16415072; GSPDB:GN00178
 A;Experimental source: strain C1p11262
 C;Genetics:
 A;Gene: pgm
 C;Superfamily: phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent
 A11750 Length: 510 May 30, 2002 09:25 Type: P Check: 2291 ..
 1 MSKSPVAII LDGFGKRAET VGNAVAQANK PNEDRYWANE PHGELKAGL
 51 DVGLPEGQMG NSEVGHNTIG AGRIVYQSLT RIDKAIEGE FOENKALNNA
 101 FTHTKENNSD LHLFGLSDG GVHSHINHLV ALLETAKDKG VKNVYIHAFL
 151 DGRDVAPOSS LEYLETLEKA ISDLNYGAIA TVSGRYAMD RDKRWERVEK
 201 AYKAIVSAEG EKFEDEPIEV KASYANDKND EFVPAITTK DGKPVATVKD
 251 NDAVIEFNFR PDRAIQLSNA FTDKEWDHFD RGHHPKNIK FVTMTLYNPS
 301 IDAEVAFEP EIMKNVIGEV L SNEGLSQLRI AETEKYPHYT FFMNGRNEE
 351 FPGENRILIN SPKVETYDK PEMSAYEYTD ALVEDIKNDK HDAIILNEAN
 401 PDMVGHGML EPTIKAIEAV DENLGRVVDL ILEKGSATIT FADHGNSETM
 451 STPEGKPHFA HTTVPVPIYV TKKGVTLEEG GRLADVAPTM LDLLGVKKPA
 501 EMTGESLIQK
 11AA_SEQUENCE 1.0
 F1;A27671 - spectrin alpha chain, nonerythroid - rat (fragment)
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 29-Sep-1999
 C;Accession: A27671
 R;Leto, T.L.; Fortugno-Erikson, D.; Barton, D.; Yang-Feng, T.L.; Francke, U.; Harris, A.S.; Morrow, J.S.; Marchesi, V.T.; Benz Jr., E.J.
 Mol. Cell. Biol. 8, 1-9, 1988
 A;Title: Comparison of nonerythroid alpha-spectrin genes reveals strict homology among diverse species.
 A;Reference number: A93097; MUID:88094373
 A;Accession: A27671
 A;Molecule type: mRNA
 A;Residues: 1-475 <LET>
 A;Cross-references: GB:M19726; NID:g205641; PIDN:AAA41678.1; PID:g205642
 A;Note: the authors translated the codon GGC for residue 18 as Ser, GAG for residue 45 as Ala, CCT for residue 46 as His, CAG for residue 48 as Pro, AGT for residue 56 as Thr, ATC for residue 61 as Ser, AGA for residue 68 as Gln, CCG for residue 75 as Leu, and GAG for residue 76 as Ala
 C;Superfamily: spectrin alpha chain; calmodulin repeat homology; SH3 homology; spectrin/dystrophin repeat homology
 C;Keywords: actin binding; EF hand
 F;205-311/Domain: spectrin/dystrophin repeat homology <SP18>
 F;320-425/Domain: spectrin/dystrophin repeat homology <SP19>

A27671 Length: 475 May 30, 2002 09:26 Type: P Check: 7377 ..
 1 RDMDEESWI KEKKLVGSE DYGRDLTGVO NLRKKHKRLE GRLAEPEQAI
 51 QGVLDGSKKL IDNTIGREE IQORPEQVE HWKELKQIAA ARGQRLSESL
 101 EYQFVANVE EEEAWINEKM TLVASEDYGD TLAAIQGLK KHEAFETDFT
 151 VHKDRVNDVC TNGQDLIKKN NHHEENISSK MKGLNGKVS LKAAQQRKA
 201 KLDENSAFLQ FNMKADVES WIGEKENSILK TDDYGRDISS VQTLTKQET
 251 FDAGLQAFQO EGIANITALK DQLLAAKHIQ SKAIEARHAS LMKRWTLQFA
 301 NSATRRKKLL EAQSHFRKVE DLFLAFAKKA SAFNSWFENA EEDLTDPVRC
 351 NSLEEIKALR EAHDAFRSSL SSAQADFNOL AELDRQIKSF RVASNPYTWE
 401 TMEALEETWR NLQKIKERE LELQKEQRRQ EENDKLROEF AQHANAFHQW
 451 IQETRTYLLD GSCIGRRVGN SGISA
 11AA_SEQUENCE 1.0
 P1;T32253 - hypothetical protein T15B7.16 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
 C;Accession: T32253
 R;Pauley, A.; Gattung, S.
 submitted to the EMBL Data Library, September 1997
 A;Description: The sequence of *C. elegans* cosmid T15B7.
 A;Reference number: 221139
 A;Accession: T32253
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-491 <PAU>
 A;Cross-references: EMBL:AF022985; PIDN:AAB69969.1; GSPDB:GN00023; CESP:T15B7.16
 A;Experimental source: strain Bristol N2; clone T15B7
 C;Genetics:
 A;Gene: CESP:T15B7.16
 A;Map position: 5
 A;Introns: 25/2; 110/3; 172/3; 304/3; 330/2; 370/2; 403/3; 432/1
 C;Superfamily: acetylcholine receptor
 T32253 Length: 491 May 30, 2002 09:26 Type: P Check: 8670 ..
 1 MEVEKQNLKR KLTQPNSPRE HSPGEPAND TISQVLRLRL FDDNYNKNTV
 51 PLKTSATVVS VEEGIONIAQ VSEISASFTL DLFSQIWHD PRLRFPHLTN
 101 CLQNLTLGIS MVEKLWTPNV CFVNSKKTETI HSSPTPIFL MIYPNGTVWV
 151 NTRLQVQSPC MYDLVLFPM D IMNCELLIES YAYNAKYKL NWREWQPVES
 201 IAKSKLSDEFT LYGLQWTKNS FEYAAGQWDQ VSSFCSLNCF SKMDGFQLTV
 251 SLTFRAYGF YILQMIPTTY SSVFLSEVSF WIDLKALPAR ITLGVSISMA
 301 LTFQYGNVAK NLPVGVYKS IDVYMLTGA FIFLTMEIVA FVCYLDSENN
 351 LRRKERQAEK KERVAVLQR KDKRKKNNYG ATTVTNANAA ESDLRSNYDE
 401 PFSQNGTTSK KISISRQANN MFESLHALAQ EGLTTDDDE NTKWTAQNVND
 451 KFCRKAFFPLS FCFNLIIYWC YLLYQNYLAK AEALAQWTPP S
 11AA_SEQUENCE 1.0
 P1;S40462 - t-complex-type molecular chaperone tcp1 (clone ASTCP-K36) - oat
 C;Species: *Avena sativa* (oat)
 C;Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 20-Aug-1999
 C;Accession: S40462; S39320
 R;Ehmann, B.; Krenz, M.; Mummert, E.; Schaefer, E.
 FEBS Lett. 336, 313-316, 1993
 A;Title: Two Tcp-1-related but highly divergent gene families exist in oat

encoding proteins of assumed chaperone function.

A:Reference number: S40461; MUID:94085629

A:Accession: S40462

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-535 <EHM>

A:Cross-references: EMBL:X75778; NID:g435174; PIDN:CA53397.1; PID:g435175

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993

C:Genetics:

A:Gene: tcp1

C:Superfamily: molecular chaperone t-complex-type

C:Keywords: molecular chaperone

S40462 Length: 535 May 30, 2002 09:26 Type: P Check: 4989 ..

```
1 MALPDEYWR PFILREQE KSRQLGLDAQ KANIAAGKSV ARLRTSLGP
51 KGMKMLQSP DGDVITINDG ATIELMDVD NQIAKIMVEL SRSQDYDIGD
101 GTTGVVVMAG SLEQAELK ERIHIPRVA EGYEMASRIA VDHLESISTK
151 YEFSATDIEP LVQTCMTTLS SKIVSRCKRA LAETAVKAVL AVADLERKDV
201 NDLIKVEGK VGGKLEDEL VQGIIVDKDM SHPQMPKRIE DAHIAILTCP
251 FEPPKPKTKH KVIDITVEKF QTLRGQEQKY FDEMVOCKKD VGATLVICOW
301 GFDEANHLH MQRELPAVRW VGGVELELIA IATGGRIVPR FQELSTEKLG
351 KAGLVREKSF GTTKDRMLYI EKCANSKAVT IFIRGNKMM IEETKRSIHD
401 ALCVARNLII NNSIYGGGS AEISCSIAVE AAADRHGVE QYAIRAFADA
451 LDATLALAE NSGLPIDTL TVVKSQHVKE NNSRCGIDCN DVGNDMKEQ
501 NVEFELIGKQ QQILATQV KMILKIDVI TPSEY
```

11AA_SEQUENCE 1.0

F1:PC2306 - dnaJ protein - *Synechococcus* sp. (strain PCC 7942) (fragment)

C:Species: *Synechococcus* sp.

C>Date: 15-Feb-1995 #sequence_revision 26-May-1995 #text_change 20-Jun-2000

C:Accession: PC2306

R:Nimura, K.; Yoshikawa, H.; Takahashi, H.

Biochem. Biophys. Res. Commun. 201, 848-854, 1994

A>Title: Sequence analysis of the third dnaK homolog gene in *Synechococcus* sp.

PCC7942.

A:Reference number: PC2306; MUID:94271242

A:Accession: PC2306

A:Molecule type: DNA

A:Residues: 1-189 <NIM>

A:Cross-references: DDBJ:D29968; NID:g559383; PIDN:BAA06235.1; PID:d1006799;

PID:g506375

C:Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

F:6-71/Domain: dnaJ amino-terminal homology <DND>

PC2306 Length: 189 May 30, 2002 09:26 Type: P Check: 7087 ..

```
1 MQNFDDYAL LGIPQADQA AIKAAFRKLA RQCHPDLPNG DRQAEEERFKQ
51 ISEAYEILSD PDRRAEQRF SRYWQOQGAA SVGSDDDDYGD FPDIEDIFVDE
101 LIGRRIVERS PRRSARRSAA TSSALSRLDE RSLEYDPKTA LQGSQAQLQL
151 EDGRLEVDI PAGIQAGEYL RLRGQGIKGG DLLLRVQLQ
```

11AA_SEQUENCE 1.0

P1:JC5550 - dnaJ protein - *Synechococcus* sp.

C:Species: *Synechococcus* sp.

C>Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 20-Jun-2000

C:Accession: JC5550

R:Oguchi, K.; Nimura, K.; Yoshikawa, H.; Takahashi, H.

Biochem. Biophys. Res. Commun. 236, 461-466, 1997

A>Title: Sequence and analysis of a dnaJ homologue gene in *Cyanobacterium*

Synechococcus sp. PCC7942.

A:Reference number: JC5550; MUID:97382457

A:Accession: JC5550

A:Molecule type: DNA

A:Residues: 1-287 <OGU>

A:Cross-references: DDBJ:AB003519; NID:g2293346; PIDN:BAA21679.1; PID:g2293347

A:Experimental source: strain PCC7942

C:Comment: This protein cooperates with DnaK protein in protein folding under various stress conditions as well as under normal conditions. In the mechanism of action of the DnaK chaperone machine, this DnaJ stimulates the hydrolysis of DnaK-bound ATP.

C:Genetics:

A:Gene: dnaJ

C:Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

F:6-71/Domain: dnaJ amino-terminal homology <DND>

JC5550 Length: 287 May 30, 2002 09:26 Type: P Check: 2256 ..

```
1 MQNFDDYAL LGIPQADQA AIKAAFRKLA RQCHPDLPNG DRQAEEERFKQ
51 ISEAYEILSD PDRRAEQRF SRYWQOQGAA SVGSDDDDYGD FPDIEDIFVDE
101 LIGRRIVERS PRRSARRSAA TSSALSRLDE RSLEYDPKTA LQGSQAQLQL
151 EDGRLEVDI PAGIQAGEYL RLRGQGIKGG DLLLRVQLQA SNFQVQGSVD
201 IYTLNVPAM AVLGQVTVP TLDGPVQMKL PASLRSGQRL RLAGKGYSKP
251 SGDRGDQIVV IQLQLPTRLS PEERQVEQL RSLQSR
```

11AA_SEQUENCE 1.0

P1:JC5648 - terminal protein precursor - *Ovine adenovirus* OAV287

C:Species: *Ovine adenovirus* OAV287

C>Date: 28-Oct-1997 #sequence_revision 28-Oct-1997 #text_change 05-Nov-1999

C:Accession: JC5648

R:Vrati, S.; Brookes, D.E.; Boyle, D.B.; Both, G.W.

Gene 177, 35-41, 1996

A>Title: Nucleotide sequence of *ovine adenovirus* tripartite leader sequence and homologues of the IVa2, DNA polymerase and terminal proteins.

A:Reference number: JC5648; MUID:97080497

A:Accession: JC5648

A:Molecule type: DNA

A:Residues: 1-641 <VRA>

A:Cross-references: GB:U31557; NID:g1117828; PIDN:AAC55958.1; PID:g1117831

C:Superfamily: adenovirus terminal protein

JC5648 Length: 641 May 30, 2002 09:26 Type: P Check: 4120 ..

```
1 MPTAGSNSSF SLFDVPGEL LDKSGVSNDE RKFKICICGIC SNKINPLFAL
51 QIHQWQLTG QSQHTLRYLRL LTLDLNRSI ARSRTHERG IKWANRFENY
101 PVTQLDLRP RGPVTSQPPF EGEPPNLLI GYFYVARVLN QYLFDRRTYS
151 NISYKLYLSP ISFERRMTWQ ILTDCSYSIN TGSYMAION VENFSQTISQ
201 IQNAVLMDRI LGSILQADMQ GFGSAISAON QNRSFOQHYQ PFSQARFVLN
251 LQERDTYLIK LICKIKKALC KFLILSNMHN SECILDLPS EFWIELEIDE
301 FSQLEIPEAE NRINKDLAT VLTIGKGMH GGALTLRSQT RVGLPRFLRP
351 RENYRAITEI MRRSRGDVIR RFIDRLPVNR RQVRVRYEET SSSPLTYPEI
401 EDEGFSDRER SSSDLSREEF NDEVIASIVD LIQNLSEELT PEARRSNFFN
451 YGSEFFQLLI RFYNENRLTD DFIQKMLVYF FILEHVA STL YLYLNLVON
501 RLARNIGIQ FVQIILRGKN ENGEDITRV WFNRQQA FR QLYSRITRDF
551 LGITEANERS YSFSTPERD QLLQDIDIVE DSGSIEEYIN QVNTDPSDL
```


601 SVEIAFRIKL SGIVGYSTNE VVLRSEFRVR EALNRMLQR Q
!!AA_SEQUENCE 1.0
PI:T08841 - polypotein - douroucouli hepatitis GB virus A
C:Species: douroucouli hepatitis GB virus A
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C:Accession: T08841
R:Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.;
Mushahwar, I.K.
J. Gen. Virol. 79, 41-45, 1998
A:Title: Genomic analysis of two GB virus A variants isolated from captive
monkeys.
A:Reference number: 216486; MUID:98120818
A:Accession: T08841
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3005 <ERK>
A:Cross-references: EMBL:AF023425; NID:g2828599; PIDN:AAC40502.1; PID:g2828600
C:Superfamily: hepatitis C virus genome polypotein
C:Keywords: polypotein

T08841 length: 3005 May 30, 2002 09:26 Type: P Check: 8233 ..

1 MEVLLVLLK TALAGAFLOP ASHACHAGY YLLTNCSDS EITACFDGCG
51 LVSVCCTVCD GRCMDLYRPG VATRPGHPGG ELFGALMSTS SPAVVAAYVA
101 GITGLGEPFS AALLAALAR VEMTPRPNT CLMDCDLAQ SEAMDIVEEM
151 RDTFWTVKRV MTLPRMLWG LVGMSFALMV VVLLLEQR VVMVLIVLAM
201 AGSVEFTGG CRCHVNGSIV AIPGAKPEDR PNATQVCVCP EGQMYLPLAL
251 CAGLAWRNGD VSGTTRDLPV RCPESMHGRG SVMCWGSAH WSWRLATEVR
301 LMEQLPGSAL CHEFAVGTTD RPHATDVLG THGIPCASCY VDRASWCGN
351 CTRDCWEKGT NKRLSEACG LGSRLTAEW AHVVDGTEG KIYDVGERP
401 RYPSHGVGT HTTAVVAARN YTVSDIGGYW HATACPNPFS PSALPKLIPG
451 RPNACLITIC KEGRLNTAW QAPGXEPPI ETECNWPRFS GVRVCEGYAF
501 DFPKSTGFI RSRAGEQQA RAGSMPPHFW LLDYLFILL VLMKLAEARL
551 VPLVAVALYW WFNMAEAT LRLHPVNI TTAAPMTWPA LPVPTVCPPT
601 KAGAVGDALY MAACLAGNAT LGVVSSVWG AYGAAEAGAR GLWRWGGFR
651 SFCAGLAWLT NVGAYLPVVE AAVGPEFVSA PLLVWAMEDS IAVFILMAN
701 IWAYWSDHLG VKLAGLVAAH LAKGALPLVL LVAAXVTRHR HSVLGFEVCY
751 SLDADASPSD WSMWALAAYV SWCLLTGLT THGRLRLE FYSYWCRRHQ
801 AVRMRLCSK VGRGRDMRV TAVWVAAGIT FPREVVRCSA ALTALALID
851 SIDYLLLETL LTAQPARAA RLDSLTFLG DADLTRAEVR RLERRGYTLF
901 QHGGQVSXGA AAILXDLGVA LEPVSVTARD CIYVRDART LACGQRYEGL
951 PVVARRGDEV LVGVFPSVRA LPPGFVPTAP VVMQRGIGF FSVVKTSMLG
1001 RDERHEGSI VVLGTSTTRS MGTGVNGVMY TTFHGSNART LAGPVGPVNC
1051 RWMSPSDVVA VYPLPSGASC LEPCKCGTQS VMCIRNDGAL CHGRLSKIVE
1101 LDLPTEISDF RSGSSGPILC DEGHVGMVM SVLHRGVKVT GVRVYKPMET
1151 LPKDSQVKSE APPVPGKTGF TEAPLYLPTG SGKSTRIPLE YTKAGHKVLY
1201 LNPSIATVRA MGPIYMEKLSG QHPSIYCGHD TTAYSRTTGS PLTYCTYGRF

1251 MANRRRYLGG ADIVICDECH VTDPTSVLGM GRARLLAREC GVRLLLFATA
1301 TPPGAPLAQH ESIEKEVPLGV DGEVAFYGHK LPVERYRTGR HLLFCHSKVE
1351 CNRLHAALST AGCNNAVYYR GNEQEIPAGD VCVCATDALS TGYTGFSFV
1401 TDCGLMVEEV VEVTLDPTIT ISVRTTPAPA ELRAQRGRG GRGSQGTYYY
1451 AMTASAPAGT LRSGLPMAAV EAGVAMYNLE PDTMADILRA YDACPYTAI
1501 TASYGEALNF FSGLVPMRNY PQVAAKSHQ HNWPLLVGVQ RTMQEANYA
1551 GPDDGPEWAG IAGTGPIPLL CRMGARPPPS VAPHHWVDL QARLGAEGY
1601 SPCYAGPILL VGLALAGAV LAHWTSGLVV VTSWRVNGNG NPLIQSTRG
1651 VSTSAPIYQLT VCVEGEQTPA DGKCAAEAVQ LXESTCGWCP MAASFDCAGM
1701 KGYLDSMRTT AAAAVEKSDS LMRSCANNY CPPGGATSA SAFASLDTK
1751 FAQAWDAIFT NGRSLVLGV AGYGARRNP LGVAAAFIMG MSAGHQVHR
1801 LAAALLLGVG GIMLGTPSVG LAMGAYFAG GSITSSLSA IVAVLGWEK
1851 AXNAASLTFD FLTGRAELKD MWFLVSCXAS PGASVAGVAL GLLLMSMKK
1901 VGEDWVNRL TLPRGSVLP DGEFVKSEFT ERVSTILRKM SLSRWMTLV
1951 ERRELDLET CSSMLMDLID MLVRFGRYIG RRLKGMVPSV RVPLVCGTIPG
2001 WGSWVGEH IEARCACGCI ITADVEEGL DVHYSSRLC SNYLGTVFY
2051 SAAGSDAEP EYPAGPALYQ IGVAEWQLV RKDKTLVGA SSVYHLHIDE
2101 LRRAIRGPPM FVGGVGSWE APLQOPPLY RAGQSVRFDD VRYSLPHTLA
2151 LPPPRPPP LAMP PPPP EQVWTEEEA DLREARARAI EAVNERLPVP
2201 NPEAAQALD ALEAAVSIL PHYGAILGDD CSCESEFGH FIPEPDAVY
2251 PIGMMEVQVG PLRDQARDLG DRLAVLGARL ESLAEHPEA SLNTERRTMG
2301 ELIDTLADVQ ARLDVTCRSD TSSSFEQIS LSDSEPETIV EGGLEVEVR
2351 PQVRFKDLI RPEGAKRLV TVRQSCADR SATRAFALSL PIAAVTATIS
2401 FDLTDHTVSD STGRVLDPLE LLQNAVGDLI VACRSESSV SYSIWSGAP
2451 LNTGRHQPAR MTRPIGTHIT ADTKVYVTD PNQAGERAAK VTIWRGSRVY
2501 DAHYRGVSE VLOQAKTIKS PGWYDEALA KVHRAAAGF GSKVTGNNIT
2551 TPARAEVNN MLAKIKTKOE VPFTLVTKRE VEFQKTRKP PRFICFPPLD
2601 FRIAEKMLG DPLVAKGIL GKSYLEFQYTP NQRYKLMVDL WRQKKHPRAI
2651 TYDATCFDSS IDERDMAVET EYFAASPNP DLVRLGSTRY AEGPMVSPKG
2701 VPLGVRKCRS SGVLTSSAN SITCYIKVKA AARAVGLVDP DFLIAGDCV
2751 IYEDDGEDH ADALVALGN XYDYCKPKTH ASLDTAESCS SYLAECNVGT
2801 ERVWMLSTDN RKPLARAASE YSDPVSSALG TILMPWHPI VRWVLLPHIL
2851 IMAFRGGGTP DDLVCEVOG NHYSPLRVL PEVLVSLHCP RCLRVATDST
2901 KTKMEAGAL RDLGMHTLAF YRKRAGNVRT RLLRGKGWG RLARALIMHP
2951 GLKEHPPSIK SIPGFKMATP YEHHTVWYS GEKPPWYMDV KCLFGLVCAF
3001 LTALL

!!AA_SEQUENCE 1.0

P1:H89844 - hypothetical protein SA0682 [Imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: H89844
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Aoki, K.; Nagai, Y.; Lian, J.; Ito, T.; Kanamori, M.; Matsumaru, H.; Maruyama, A.; Murakami, H.; Hosoyama, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; Hirakawa, H.; Kihara, S.; Goto, S.; Yabuzaki, J.; Kanehisa, M.; Yamashita, A.; Oshima, K.; Furuya, K.; Yoshino, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H89844
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-501 <KUR>
A:Cross-references: GB:BA000018; PID:g13700618; PIDN:BAB41915.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA0682
C:Superfamily: peptide transporter protein
H89844 Length: 501 May 30, 2002 09:26 Type: P Check: 6701 ..
1 MTQONSHGNQ IODIPOTGF GHPRGLGLF EWEFERFSY YGMRALLIYF
51 MYFAVTDNGL GIDKTTAMSI MSVYGLIYM TSIPGWIMD RITGRTGATL
101 LGAVFTIIGH ICLSLPFLI GLFTSMFII IGSGLMKPNI SNIVGRLYPE
151 NDRRMDAGFV IFYMSVNMGA LLSPIILQHF VNVKNFHGF LIAAVGMALG
201 LVWTVLFNRK NLGSVGMKPT NPLTPAEKKK YGLIGSVVL AIVLIIVIGA
251 LTNLSLSENLV SNTVLVLGIA LPIIYFTLLI RSKDVTDTTER SRVKAFLPLF
301 ILGMVFMAIQ EOGSNVLNIY GIEHSDMKLN LEGWKTNFGF AIFQSINPLF
351 ILLAPITISL LMQKLGTKQP SLPVKFAIGT FLAGASYILI GIVGYASGSS
401 NFSVNMVILS YIICVIGELC LSPTGNSAAV KLAPKAFNAQ MMSIWYLTNA
451 SAQAINGTLV KLIEPLGQTN VFIFLGVVAI IVTTIVLAFS PLIIKAMKGI
501 R
!!AA_SEQUENCE 1.0
P1:C97003 - ABC-type spermidine/putrescine transport system, permease component II CAC0838 [Imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: C97003
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, H.M.; Dubois, J.; Oiu, D.; Hitti, J.; Wolf, Y.I.; Tatusov, R.L.; Sabathe, F.; Doucette-Stamm, L.; Soucaille, P.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum.
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C97003
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78814.1; PID:g15023730; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0838
C:Superfamily: spermidine/putrescine transport system permease protein potI

C97003 Length: 260 May 30, 2002 09:26 Type: P Check: 8409 ..
1 MVEKWLKRFY LFTIYIFLYA PIVELMVFSF NSEKSSHWG HESLTWYKAL
51 LODDRILLTAL YVTVLVAIVS SIISTIFGTI SAIGISKMSF LPKLLNLVN
101 NIPVLNPDIV MAVSLMTLFI FLKIPGLTT LIAHIAFSV PYVILSVLPK
151 LTQLPTDIVK AALDGGATPS YAMRKILLPQ IKSCIIAGFL FAFMTSIDDF
201 VISFNTGND VTNLSIEIYS MARGITPEI NALSTLMFVT ILLLLANR
251 KSISKEKK
!!AA_SEQUENCE 1.0
P1:C84811 - phosphate transporter (AtPT2) [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: C84811
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: C84811
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-534 <STO>
A:Cross-references: GB:AE002093; NID:g3928081; PIDN:AACT9607.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g38940
A:Map position: 2
C:Superfamily: probable inorganic phosphate transport protein PHO84
C84811 Length: 534 May 30, 2002 09:26 Type: P Check: 605 ..
1 MAREQLOVLN ALDVAKTQWY HFTALIIAGM GFTIDAYDLF CISLVTKLGG
51 RIYHVEGAQ KPQTELPNVA AAVNGVAFCG TLAGOLFFGW LGDKLGRKKV
101 YGMTLWVWL CSIASGLSFG HEPRKAVMATL CFFRWLGFG ICGDYPISAT
151 IMSEYANKKT RGAIFYSAVFA MÖFGIMAGG IFALIISSAF EAKFPSPAYA
201 DDALGSTIPQ ADLVRIILM AGAIPAMTY YSRKMPETA RYALVAKDA
251 KQASDMSKV LÖVEIEPEQÖ KLEISKEKS KAFGLFSKEF MSRHGLHLGG
301 TTSTWFLIDI AFYSQMLFÖK DIFSAIGWIP PAÖSMNAIÖE VFKIARAÖTL
351 IALCSTVPGY WFTVAFIDVI GRFAIÖMMGF FEMTVEMFAL AIPYNHWTÖK
401 ENRIGEVIMY SLTFPEANFG PNAITPVVPA EIPPARFRST CHGISAASGK
451 LGAMVGAFGF LYLAONPDKD KTDAGYPPGI GVRNSLIVLG VVNFGLILFT
501 FLVPESKGS LEEMSGENED NENSNDSTRT VPVIV
!!AA_SEQUENCE 1.0
P1:T07607 - phosphate transport protein 2 - potato
N:Alternate names: Inorganic phosphate transporter 2
C:Species: Solanum tuberosum (potato)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C:Accession: T07607
R:Leggewie, G.; Willmitzer, L.; Riesmeier, J.W.
Plant Cell 9, 381-392, 1997
A:Title: Two cDNAs from potato are able to complement a phosphate

uptake-deficient yeast mutant: Identification of phosphate transporters from higher plants.

A:Reference number: Z07579; MUID:97246321

A:Accession: T07607

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-527 <LEG>

A:Cross-references: EMBL:X98891; NID:g1420872; PIDN:CAA67396.1; PID:g1420873

A:Experimental source: cv. Desiree

C:Comment: Expressed mainly in root organs when plants are deprived of phosphate and to a lesser extend under sulfur deprivation conditions.

C:Genetics:

A:Gene: PT2

C:Superfamily: probable inorganic phosphate transport protein PH084

C:Keywords: membrane protein; phosphate transport

T07607 Length: 527 May 30, 2002 09:26 Type: P Check: 2263 ..

1 MAVEDNNLQV LNALDLAKTQ LYHFTAIITA GMGFTDAYD LFSISLVTKL

51 LGRLYYTKPD LKPGTLPPA RVGLRHGAL VGTLAGQLFF GCARLAKMGR

101 KKVYGMTLVL MIVCSVASGL SLGNTPKVVM TTLCFERFWM GFGIGDYPPL

151 SATIMSEYAN KTRGAFIAA VFAMQFGIL FSGIVALITA AGFDHAYKAP

201 TPAENAAVST VPQADYIWR ILMFGLPAA LTYYWRMKMP ETARYTALVA

251 KDAKRAQDM GRVLQVEIES EEAKIEQISR DETNQFGLFS WEFVRRHGLH

301 LFGTCTSWFL LDIAFYSONL FQKDVFSAVG WIPKAPTUNA VOELYKTIARA

351 QTLIALCSTV PGWFTVAFI DIIGRAIQL MGEFFMTVFM FAIAIPYHHW

401 TLEANRIGFI VMYSLTFEFA NEGPNAITFV VPAEIFPARL RSTCHGISAA

451 AGKAGAIVGA YGFLYAQSK DPMKTDAGYP AGIGIKNSLI VLGFINALGM

501 VCTFCVPESK GKSLEASOE TISTGEA

!!AA_SEQUENCE 1.0

P1:T07604 - phosphate transport protein PT1 - potato

N:Alternate names: inorganic phosphate transporter 1

C:Species: Solanum tuberosum (potato)

C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000

C:Accession: T07604

R:Leggewie, G.; Willmitzer, L.; Riesmeier, J.W.

Plant Cell 9, 381-392, 1997

A:Title: Two cDNAs from potato are able to complement a phosphate uptake-deficient yeast mutant: Identification of phosphate transporters from higher plants.

A:Reference number: Z07579; MUID:97246321

A:Accession: T07604

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-540 <LEG>

A:Cross-references: EMBL:X98890; NID:g1420870; PIDN:CAA67395.1; PID:g1420871

A:Experimental source: cv. Desiree; root

C:Genetics:

A:Gene: PT1

A>Note: expressed in roots, tubers and source leaves as well as in floral organs

C:Superfamily: probable inorganic phosphate transport protein PH084

C:Keywords: phosphate transport

T07604 Length: 540 May 30, 2002 09:26 Type: P Check: 8926 ..

1 MANDLQVLNA LDVAKTQLYH FTAIVAGMG FETDAYDLFC ISMVKLLGR

51 IYYHHDNALK PGSLLPNVSA AVNGVAFCGT LAGQLFFGWL GDKMGRKKVY

101 GMTLMIMVIC SIASGLSFCH TPKSVMTTLC FFRFWLGFGL GGDYPLSATI

151 MSEYANKTR GAFIAAVFAM QGFGLAGGM VAIYSSAFK GAFPAAYEV

201 DALASTVSQA DEFWRIILMF GAIPAGLTYV WRMKPETAR YTALVAKNLK

251 QAANDMSKVL QVEIEAPEK VAAISVANGA NEFGLFSKEF LRRHGLHLIG

301 TASTWFLLDI AFYSQNLFOK DIFSAIGWIP PAOTMNALEE YKIKARAQTL

351 IALCSTVPGY WFTVAFIDRI GRAFIQLMGF FFMVFMFAL ALPHYHWTLK

401 DNRIGFVVMY SLTFFFANFG PNATTFVPA EIFPARLRST CHGISAAAGK

451 AGAMVGAFFG LYAAQPTDPK KTDAGYPAGI GVRNSLIYLG CVNFGMLFT

501 FLVPESKGS LEEMSRENEG EETVAEMRA TSGRTVLFFK

!!AA_SEQUENCE 1.0

P1:T07740 - probable inorganic phosphate transport protein 1 - tomato

C:Species: Lycopersicon esculentum (tomato)

C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000

C:Accession: T07740

R:Daram, P.; Brunner, S.; Persson, B.L.; Amrhein, N.; Bucher, M.

Planta 206, 225-233, 1998

A:Title: Functional analysis and cell-specific expression of a phosphate transporter from tomato.

A:Reference number: Z16107; MUID:98408214

A:Accession: T07740

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-538 <DAR>

A:Cross-references: EMBL:Y14214; NID:g2266671; PIDN:CAA74607.1; PID:g2266672

A:Experimental source: cultivar MoneyMaker; root hair

C:Genetics:

A:Gene: LePT1

C:Superfamily: probable inorganic phosphate transport protein PH084

C:Keywords: transmembrane protein

T07740 Length: 538 May 30, 2002 09:26 Type: P Check: 4193 ..

1 MANDLQVLNA LDVAKTQLYH FTAIVAGMG FETDAYDLFC ISMVKLLGR

51 IYYHHDGALK PGSLLPNVSA AVNGVAFCGT LAGQLFFGWL GDKMGRKKVY

101 GMTLMIMVIC SIASGLSFCH TPKGVMTTLC FFRFWLGFGL GGDYPLSATI

151 MSEYANKTR GAFIAAVFAM QGFGLAGGM VAIIVSAERK GAFPAAYEV

201 DAIGSTVPOA DEFWRIILMF GAIPAGLTYV WRMKPETAR YTALVAKNLK

251 QAANDMSKVL QVEIEAPEK VTAISEAKGA NDFGLFTKEF LRRHGLHLIG

301 TASTWFLLDI AFYSQNLFOK DIFSAIGWIP PAOTMNALEE YKIKARAQTL

351 IALCSTVPGY WFTVAFIDKI GRAFIQLMGF FFMVFMFAL AIPYHWTLK

401 DHRIGFVVMY STFFFFANFG PNATTFVPA EIFPARLRST CHGISAAAGK

451 AGAMVGAFFG LYAAQPTDPT KTDAGYPPGH WCEELVDRPW LCNFLGMLFT

501 FLVPESNGKS LEDLSRENEG EETVAEIRA TSGRTVPV

!!AA_SEQUENCE 1.0

P1:T01124 - probable phosphate transporter At2g32830 [imported] - Arabidopsis

thaliana

N:Alternate names: hypothetical protein F24L7.3

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Feb-2001

C:Accession: T01124; T00785; A84738

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,

S.M.; Kaul, S.; Mason, T.M.; Kerlavage, A.R.; Adams, M.D.; Somerville, C.R.;

Venter, J.C.

submitted to the EMBL Data Library, December 1997

A:Description: Arabidopsis thaliana chromosome II BAC T21L14 genomic sequence.

A:Reference number: Z14209

A:Accession: T01124
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-542 <ROW>
A:Cross-references: EMBL:AC003033; NID:g2702261; PID:g2702279
A:Experimental source: cultivar Columbia
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, T.M.; Kerlavage, A.R.; Adams, M.D.; Somerville, C.R.; Venter, J.C.
submitted to the EMBL Data Library, February 1998
A:Description: Arabidopsis thaliana chromosome II BAC F24L7 genomic sequence.
A:Reference number: Z14204
A:Accession: T00785
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-542 <ROW>
A:Cross-references: EMBL:AC003974; NID:g2914688; PID:g2914691
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Cressy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: A84738
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-542 <STO>
A:Cross-references: GB:AE002093; NID:g2702279; PID:AAB91982.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g32830
A:Map position: 2
A:Introns: 102/1; 259/3
A:Note: T21L14.23; F24L7.3
C:Superfamily: probable inorganic phosphate transport protein PH084

T01124 Length: 542 May 30, 2002 09:26 Type: P Check: 992 ..

1 MARKGKVLN ALDAKTOY HFTALVIAGM GFETDAYDLF SISLVTKLG
51 RIYVHVDSK KBGTLPPNVA AAVNGVAFCG TLAGLFFGW LGDKLGRKKV
101 YGITLMLMV LCSIQSGLSFG HSANGVMATL CFFRFLGFG IGGDYPLSAT
151 IMSEYANKKT RGAFTAAVFA MGFEGILAGG IVSLIVSSTF DHAFAKPTYE
201 VDPVGSTVPQ ADVVWRIVLM FGAIPALLTY YWRMKMPETA RYTALVARNT
251 KQASDMSKV LQVDLIAEE AQSNSNSNP NFFGLFTRE FARRHGLHL
301 GTTTFWLLD IAVYSSNLFQ KDITTAIGWI PAETMNAIH EYFTVSKAQT
351 LIALCGTVPG YWFTVAFIDI LGRFTIQLMG FIFMTIFMFA LAIPYDHWRH
401 RENRIGFLIM YSLTMEFANF GPNATTFVVP AEIFPARLRS TCHGISAASG
451 KAGAIVGAFG FLYAAQSSDS EKTDAGYPPG IGVNRSLLML ACVNFLGIIVF
501 TLLVPESKKG SLEEISREDE EDSGCDTVE MTVANSGRKV PV

11AA_SEQUENCE 1.0
P1:T07892 - probable inorganic phosphate transport protein PT1 - barrel medic
C:Species: Medicago truncatula (barrel medic)
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 23-Mar-2001
C:Accession: T07892
R:Lin, H.; Trieu, A.T.; Blaylock, L.A.; Harrison, M.J.
Mol. Plant Microbe Interact. 11, 14-22, 1998
A:Title: Cloning and characterization of two phosphate transporters from

Medicago truncatula roots: regulation in response to phosphate and to
colonization by arbuscular mycorrhizal (AM) fungi.
A:Reference number: Z16196; MUID:98086876
A:Accession: T07892
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-537 <LIU>
A:Cross-references: EMBL:AF000354; NID:g2465526; PID:AAB81346.1; PID:g2465527
C:Genetics:
A:Gene: PT1
C:Superfamily: probable inorganic phosphate transport protein PH084

T07892 Length: 537 May 30, 2002 09:26 Type: P Check: 33 ..

1 MSGELGVNA LDVAKTQLYH FTTIVIAGMG FETDAYDLFC ISLVTKLIGR
51 IYTEDNPTR PGTLPSSAQ S AVTGVALLVT LAGQLFFGWL GDKLGRKKV
101 GLTILMVVC SVGSLSFGS SPKSVMATLC FFRFWLFGFI GGDYPLSATI
151 MSEYANKKTR GAFTAAVFAM QGFGLGGI VALIVASIFD HKYKVPTFEE
201 NPATSLVLPQ FDYVWRILIM FGAIPALTY YWRMKMPETA RYTALVAKNA
251 KQAAADMSKV LQVELEVEEE KYQKMTSDKR NSYGLFSKQF AARHGLALFG
301 TCSTWFLIDI AFYSQNLFOK DIFSAIGWIP PAKEMNAIHE YKTIARAQTL
351 IALCSTVPGY WFTVAFIDHM GRFAIQMMGF FEMTVEMFGL AIPYDHSKE
401 ENRIGFVVMY SLTFFSNFG PNAATFVPA EIFPARLRST CHGISAAAGK
451 AGAIVGAFGF LYAOSKDPPT KTDKGYPTGI GINSLIMLG VINFGMLCT
501 LLVPESKGS LLELSENEG EGAEATEQEG PREENVA

11AA_SEQUENCE 1.0
P1:T07894 - probable inorganic phosphate transport protein PT2 - barrel medic
C:Species: Medicago truncatula (barrel medic)
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 23-Mar-2001
C:Accession: T07894
R:Lin, H.; Trieu, A.T.; Blaylock, L.A.; Harrison, M.J.
Mol. Plant Microbe Interact. 11, 14-22, 1998
A:Title: Cloning and characterization of two phosphate transporters from
Medicago truncatula roots: regulation in response to phosphate and to
colonization by arbuscular mycorrhizal (AM) fungi.
A:Reference number: Z16196; MUID:98086876
A:Accession: T07894
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-533 <LIU>
A:Cross-references: EMBL:AF000355; NID:g2465528; PID:AAB81347.1; PID:g2465529
C:Genetics:
A:Gene: PT2
C:Superfamily: probable inorganic phosphate transport protein PH084

T07894 Length: 533 May 30, 2002 09:26 Type: P Check: 2173 ..

1 MSGELGVNA LDVAKTQLYH FTTIVIAGMG FETDAYDLFC ISLVTKLIGR
51 IYTEDNPTR PGTLPSSAQ S AVTGVALLVT LAGQLFFGWL GDKLGRKKV
101 GLTILMVVC SVGSLSFGS SPKSVMATLC FFRFWLFGFI GGDYPLSATI
151 MSEYANKKTR GAFTAAVFAM QGFGLGGI VALIVASIFD HKYKVPTFEE
201 NPASLLVLPQ FDYVWRILIM FGAIPALTY YWRMKMPETA RYTALVAKNA
251 KQAAADMSKV LQVELEVEEE KYQKMTSDKR NSYGLFSKQF AARHGLALFG
301 TCSTWFLIDI AFYSQNLFOK DIFSAIGWIP PAKEMNAIHE YKTIARAQTL
351 IALCSTVPGY WFTVAFIDHM GRFAIQMMGF FEMTVEMFAL AIPYDHSKE

401 ENRIGFVVIY SLTEFFANFG PNATTFVVP A EIFPARLRST CHGISAAAGK
451 AGAIVGAFGF LYAAQSKDPT KTDKGYPTGI GIKNSLIMLG VINFGMLCT
501 LLVPESNGKS LEELSGENEG EGAEATEQEG SRV

!!AA_SEQUENCE 1.0
P1:T05714 - probable inorganic phosphate transport protein - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 08-Oct-1999
C:Accession: T05714
R:Liu, C.; Muchhal, U.S.; Mukatira, U.; Kononowicz, A.K.; Raghothama, K.G.
A:Description: Tomato phosphate transporter genes are differentially regulated
in plant tissues by phosphorus.
A:Reference number: Z15426
A:Accession: T05714
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-538 <LIU>
A:Cross-references: EMBL:AF022873; NID:g2570522; PIDN:AAB82146.1; PID:g2570523
C:Genetics:
A:Gene: PT1
C:Superfamily: probable inorganic phosphate transport protein PH084

T05714 Length: 538 May 30, 2002 09:26 Type: P Check: 4793 ..

1 MANDLOVLNA LDVAKTQLYH FTAIVIAGMG FETDAYDLFC ISMVTKLIGR
51 LYVHHDGALK PGSLLPNVSA AVNGVAFCGT LAGQLFFGWL GDKMKRKKVY
101 GMTLMIMVIC SIASGLSFGH TPKGVMITLC FFRFWLGFGI GGDYPLSATI
151 MSEYANKKTR GAFIAAVFAM QGFGLAGM VAIIVSAAFK GAFAPAPAYEV
201 DAIGSTVPQA DEWVRILMF GAIPAGLTYG WRMKMPETAR YTALVAKNLK
251 QAANDMSKVL QVEIEAPEK VTAISEAKG NDFGLFTKEF LRRHGLHLG
301 TASTWFLLDI AFYSQNLFOK DIFSAIGWIP PAQTMNALLEE YKIKARAQTL
351 IALCSTVPGY WETVAFIDKI GRFAIQLMGF FEMTFMEAL AIPYHWTLK
401 DHRIGFVVMY SFTFFANFG PNATTFVVP A EIFPARLRST CHGISAAAGK
451 AGAMVGAFGF LYAAQPTDPT KTDAGYPPGI GVRNSLIYLG CVNFGMLEFT
501 FLVPESNGKS LEDLSRENEG EETVAEIRA TSGRTVPV

!!AA_SEQUENCE 1.0
P1:T07164 - probable inorganic phosphate transport protein - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Jun-2000
C:Accession: T07164
R:Tahiri-Alaoui, A.; Avrova, A.; Antoniw, J.F.
A:Description: A phosphate transporter from tomato roots up-regulated during
arbuscular mycorrhiza colonization.
A:Reference number: Z15971
A:Accession: T07164
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-538 <TAH>
A:Cross-references: EMBL:Y16125; PIDN:CAA76075.1
A:Experimental source: cultivar Early Mech; root
C:Superfamily: probable inorganic phosphate transport protein PH084

T07164 Length: 538 May 30, 2002 09:26 Type: P Check: 5788 ..

1 MANDLOVLNA LYVAKTQLYH FTAIVIAGMG FETDAYDLFC ISMVTKLIGR
51 LYVHHDGALK PGSLLPNVSA AVNGVAFCGT LAGQLFFGWL GDKMKRKKVY

101 GMTLMIMVIC SIASGLSFGH TPKGVMITLC FFRFWLGFGI GGDYPLSATI
151 MSEYANKKTR GAFIAAVFAM QGFGLAGM VAIIVSAAFK GAFAPAPAYEV
201 DAIGSTVPQA DEWVRILMF GAIPAGLTYG WRMKMPETAR YTALVAKNLK
251 QAANDMSKVL QVEIEAPEK VTAILKQKA NDFGLFTKEF LRRHGLHLG
301 TASTWFLLDI AFYSQNLFOK DIFSAIGWIP PAQTMNALLEE YKIKARAQTL
351 IALCSTVPGY WETVAFIDKI GRFAIQLMGF FEMTFMEAL AIPYHWTLK
401 DHRIGFVVMY SFTFFANFG PNATTFVVP A EIFPARLRST CHGISAAAGK
451 AGAMVGAFGF LYAAQPTDPT KTDAGYPPGI GVRNSLIYLG CVNFGMLEFT
501 FLVPESNGKS LEDLSRENEG EETVAEIRA TSGRTVPV

!!AA_SEQUENCE 1.0
P1:T05724 - probable inorganic phosphate transport protein PT2 - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 08-Oct-1999
C:Accession: T05724
R:Liu, C.; Muchhal, U.S.; Mukatira, U.; Kononowicz, A.K.; Raghothama, K.G.
A:Description: Tomato phosphate transporter genes are differentially regulated
in plant tissues by phosphorus.
A:Reference number: Z15426
A:Accession: T05724
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-528 <LIU>
A:Cross-references: EMBL:AF022874; NID:g2570524; PIDN:AAB82147.1; PID:g2570525
C:Genetics:
A:Gene: PT2
C:Superfamily: probable inorganic phosphate transport protein PH084

T05724 Length: 528 May 30, 2002 09:26 Type: P Check: 3242 ..

1 MAVGDNDNNN LQVLNALDIA KTQLYHFTAI VIAGMGFTD AYDLFSISLV
51 TKLLGRLYYT KPDLLKPGTL PPAVSASVVG VALVGTLAGQ LFFGWLGDKM
101 GRKKVYGMTL VLMVVCVSAV CLSEGSTPKG VMTTLCFFRF WLFGIGGDY
151 PLSATIMSEY ANKTRGAFI AAVFAMQFG ILFSGIVALI TAAGFDHAYR
201 SPTEENNAL STVQSDIYI RIILMFGSLP AALTYWRMK MPETARYTAL
251 VAKDAKRAAQ DMGKVLQVEI ESEAKIEOI SRNETNOFGL FSWEFVRRHG
301 LHLEFGTSTW FLIDIAFYSQ NLFQKDVESA VGMIPKAPTM NAVQEVYKIA
351 RAQTLIALCS TYPGYWFTVA FIDIIGRFAT QLMGFEFTV FMAIATPYH
401 HMTLEANRIG FIVMYSLTFE FANFGPNATP FVVPALIFPA RLIRSTCHGIS
451 AAAGKAGATV GAYGFLYAAQ SKDPNKTDAQ YPAGIGIKNS LIVLGCINAL
501 GMLCTFCVPE PKGKSLEAS QETITGEA

!!AA_SEQUENCE 1.0
P1:T07808 - probable inorganic phosphate transport protein - Madagascar
periwinkle
C:Species: Catharanthus roseus (Madagascar periwinkle)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
C:Accession: T07808
R:Kai, M.; Masuda, Y.; Kikuchi, Y.; Osaki, M.; Tadano, T.
A:Title: Isolation and characterization of a cDNA from Catharanthus roseus
which is highly homologous with phosphate transporter.
A:Reference number: Z07788

A:Accession: T07808
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-542 <KA1>
A:Cross-references: EMBL:AB004809
C:Genetics:
A:Gene: PIT1
C:Superfamily: probable inorganic phosphate transport protein PH084

T07808 Length: 542 May 30, 2002 09:26 Type: P Check: 2442 ..

```
1 MAKEQLQVLN ALDVAKTQWY HPTAIVIAM GFFTDAYDLF CISLVTKLIG
51 RIYYHVDGAV KPGTLPPNVS AAVNGVAFCG TLAGQLFFGW LGDKMGRKRV
101 YGWTMLMVI ASIASGLSFG DKPKAVMATL CFFRFWLFGF IGGDYPLSAT
151 IMSEYANKKT RGAFLAAVFA MGFEGILAGG MVAIIVSASF KAGFPAPAYQ
201 DGAVASTVPE ADVVWRITLM FGAIPALTY YWRMKMPETA RYTALVAKNA
251 KOANDMSKV LQVELEAEQ EYKFAQEP A NTFGLTFKEF LKRHGLHLIG
301 TATTFWFLDI AFYSQNLFOK DIFSAIGWIP PAQTMAIEE VFRIARAQTL
351 IALCSTVPGY WFTVFLIDRI GPFIIQWGF FFWTFMFAL AIPYNHWTNK
401 DNRIGFVIMY SLTFFEFANFG PNATTFVPA EIFPARLRST CHGISAAAGK
451 AGAIIAGFGF LYAAQPSDPS KTDKGYRPGI GVKNALIVLG CVNPLGMVFT
501 FLVPEAKGKS LEEVSKENEE EVENGTCLRQ QSGHDTRTVP VL
```

IIAA_SEQUENCE 1.0
P1:T47629 - phosphate transport protein - Arabidopsis thaliana
N:Alternate names: protein T5N23.60
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
C:Accession: T47629
R:Obermaier, B.; Ottenwelder, B.; Duchemin, D.; Zeltler, K.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.
Submitted to the Protein Sequence Database, March 2000
A:Reference number: 224463
A:Accession: T47629
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-535 <OBE>
A:Cross-references: EMBL:AL138650
A:Experimental source: cultivar Columbia; BAC clone T5N23
C:Genetics:
A:Map position: 3
A:Note: T5N23.60
C:Superfamily: probable inorganic phosphate transport protein PH084

T47629 Length: 535 May 30, 2002 09:26 Type: P Check: 3263 ..

```
1 MAGDQLNLN ALDVAKTQWY HPTAIIAGM GFFTDAYDLF CISLVTKLIG
51 RIYYHVDGSE KPGTLPPNVS AAVNGVAFCG TLAGQLFFGW LGDKLGRKKV
101 YGWTMLVMVL CSIASGLSFG SNPKVTMTTL CFFRFWLFGF IGGDYPLSAT
151 IMSEYANKKT RGAFLAAVFA MGFEGILTGG IFALIISAAF EAKFPAPTYQ
201 IDALASTVPQ ADVVWRITLM VGALPAAMTY YSRSKMPETA RYTALVAKDA
251 KLAASNMSKV LQVEIEAEQ GTEDKSNSFG LFSKEFMKRH GLHLGTST
301 WFLDLIAFYS QNLFQKDIFS AIGWIPPAQT MNAIQEVFKI ARAQTLIALC
351 STVPGWFTV AFIDVIGRFA IQMNGFFMT VFMFALAIPT DHMTHKENRI
```

```
401 GFVAMYSLTF FFANCPNAT TEVVPALIEP ARPRSTCHGI SAASGKLGM
451 VGAFGLYLIA QSPDKTTEH GYPRGIGVKN SLIVGVNL LGMVFTLLVP
501 ESKGSLEEM SGENQDNDES SSSSNNSNN AVSTA
```

IIAA_SEQUENCE 1.0
P1:T34457 - hypothetical protein T19H12.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C:Accession: T34457
R:Davidson, S.
Submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid T19H12.
A:Reference number: 221528
A:Accession: T34457
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-533 <DAV>
A:Cross-references: EMBL:U97009; PIDN:AAC69033.1; GSPDB:GN00023; CESP:T19H12.10
A:Experimental source: strain Bristol N2; clone T19H12
C:Genetics:
A:Gene: CESP:T19H12.10
A:Map position: 5
A:Introns: 50/3; 73/3; 126/1; 156/1; 475/2
C:Superfamily: glucuronosyltransferase

T34457 Length: 533 May 30, 2002 09:26 Type: P Check: 9544 ..

```
1 MRLTFSCIC FLKRVYAAK ILIFNPIFGF SHVKEVSKLA DIADHGEV
51 TLFPPEHLAL KNLDGIVKNK NIEIINYRDP HYDDLKLET QTFPFDWDSQ
101 LMNPNYLMF MLPRILGGEF KKTIIQLIKD KEILKKLKD KEDVAISETF
151 ELTGMYMSHF LGVPCIPILS AVRLDIFNEA FGSSAFGYL TQGSKLAPD
201 AGFLRLNDV YRDEFSKMAF RGMAYQNDV IEKAAGHVP YMKDLVKEAP
251 VYMTNSNPYL DFAVPTTATI VHIGGITINL EKMNHVDALP EEEYIILKEK
301 ETTVLISFGS VIRSEMPEN FKAGLIKVFE SLDPVIFIWK YEIDDLLEFQK
351 KLPKNVHLKK WVPQPSLLAD KRVKLEVTG GLGSTMEVAY TCKPALMVPPI
401 FGDQPMNADM LARHGALAY DKFDLVDGKK LTFEVRDLVT NPKYEQKAKE
451 LLDVLTNQPI DPMVNLMKHL EFAIKFPNLR SQIPEINQV PIAYHYLDVI
501 VFLIFVSIIT AVISQIVCR ILSRLSKKV KSD
```

IIAA_SEQUENCE 1.0
P1:S56753 - interferon regulatory factor 3 - chicken
C:Species: Gallus gallus (chicken)
C>Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
C:Accession: S56753
R:Grant, C.E.; Vasa, M.Z.; Deeley, R.G.
Nucleic Acids Res. 23, 2137-2146, 1995
A:Title: cIRF-3, a new member of the interferon regulatory factor (IRF) family that is rapidly and transiently induced by dsRNA.
A:Reference number: S56753; MUID:95334365
A:Accession: S56753
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-491 <GRA>
A:Cross-references: EMBL:U20338; NID:g790580; PIDN:AAA86995.1; PID:g790581
C:Superfamily: lymphoid-specific interferon regulatory factor

S56753 Length: 491 May 30, 2002 09:26 Type: P Check: 3644 ..

```
1 MAALDSEGDA QKLRGPWLL NAVSSGLYRG LCWIDPDRRI FRIPWKHNR
51 KDVTSSDVEI FKAWAKASGR YEGNAEDPAK WKTNFRCALR STHMFMLLLED
```

101 RSKCNDPHK VYAVASGVPN DRSGGPPVAG ALQQPQLLL NHHDLALENT
151 PTDSTEGVAA AALTQVDLDL LQSVLQHCNI SALGSQPTLW AHTGDALPED
201 ALLPQGDCG LPPQFQDWR QLEEPLLIGN QPLTGGCGQ DGAGALPASE
251 ECAIPAPSPA EELLFQSANP APPPAGDIG GLPPLDITI YYRGKMVQE
301 QVDSRCVLA YQPLDPAVAE QRLVLFPSPA SLPPRQRRY TEDLLEVAGL
351 RLEQAGQLL ATRLKCKVF WALSQLEGG EPPLNLHRD QETTIFDFRY
401 FCTELRDFRD SRRERSPDFT IFLCFGQCFSS STKPESKLI LKLVPPQCE
451 YWYEQVQRCG ASSLSGNVS LQLSDSFNLF ELIEQYHMQT D

!!AA_SEQUENCE 1.0

PI:S73432 - MG096 homolog D09_crf518 - Mycoplasma pneumoniae (strain ATCC 29342)

C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342

C>Date: 26-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999

C:Accession: S73432

R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.

A:Reference number: S73327; MUID:97105885

A:Accession: S73432

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-518 <HIM>

A:Cross-references: EMBL:AE000013; GB:U00089; NID:g1673762; PIDN:AAB95754.1; PID:g1673763

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: hypothetical protein MG096

S73432 Length: 518 May 30, 2002 09:26 Type: P Check: 4533 ..

1 MNELQTKVNE ANNIFPVEAF KVPKVPKELF GFVNQGFEPK LNPGLNIAD
51 NVASLFEQYS LKQASLKPED ILLEKKNDIV LEHKVRYNFA LQHFETTYV
101 GTGGEINLOF ALQASTTNS SLEELQASFS KTGDNLTAQL FWKPTVTCLV
151 SGENDLTHIA QTAIGESLFD SRVDLSASII NSEATLKTAE ATFTQVLYNP
201 FKAEREKALA IKKAEERIK KELEEQKRRQ EELSKQORDK EALQKSIMNF
251 QEFISYWTGQ GKDVQKQKQF IQALEAASFST NMNEVENILLI AGFRSAIQTY
301 YKDGKADQSQ NAKIAFGEKG IQPPKSGPGL DGIFMSDFLR GNLIGNAHFD
351 LKLKKEVYKN TQCKDAQGND KKASINQAK QNNPFRQVN PMDFSEVEL
401 KYEGSYGLYP GARFLNFGS LGIPNDWKE MSVKFVLDGK TPQWIADKPD
451 YPGSLFEFEK NQLKFTPHVK EHVHVENKQF MEKLKSONLH NLELATGATK
501 PPVVDLASYL HYLILNHK

!!AA_SEQUENCE 1.0

PI:S73923 - MG288 homolog G07_orf417 - Mycoplasma pneumoniae (strain ATCC 29342)

C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342

C>Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999

C:Accession: S73923

R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma

pneumoniae.

A:Reference number: S73327; MUID:97105885

A:Accession: S73923

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-417 <HIM>

A:Cross-references: EMBL:AE000058; GB:U00089; NID:g1674291; PIDN:AAB96245.1; PID:g1674299

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: hypothetical protein MG096

S73923 Length: 417 May 30, 2002 09:26 Type: P Check: 7095 ..

1 MGFHAVYQNN DTKANLDFAL NISTINFATL QELQNSFDLQ GSDLTAGLFY
51 KYSVNKLTSQ TNDLTTIAKT ALGENTIQKQ VSLTQSIKP RLEAKTQYK
101 QDIAPFAKE RQALAQHLK EIEAKQRAE QLKEQQAEE KRQEEVKVY
151 AETQOFNDL TSAQKFEYV LKQKDVTKR VELIQALKSS FERNQNRFFN
201 FLIAGERTAI DWYNOEKNN TTAKNNAFGK NGIQFPVAGF QGITYMSQWLR
251 DELSGKTDIK LNLKSLSVQ ENKNSSINWN KQRIEIKQV KPFNYSFEIN
301 LKVTGSYNVS LWYLIGAAGT GIPTWSGTM DMKFIVDGL DSGIVTKQDY
351 PGSKEFTED KLMFTLHVQK QIKVKEQGFV NLLKGQSLDN LDLRTGTTRP
401 PYVDLASYLH FVILTAK

!!AA_SEQUENCE 1.0

PI:B72719 - probable bacterioferritin comigratory protein APE0291 - Aeropyrum

pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: B72719

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takahashi, M.; Sekine, M.; Baba, S.; Ankai, A.; Kosugi, H.; Hosoyama, A.; Fukui, S.; Nagai, Y.; Nishijima, K.; Nakazawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, T.; Kushiida, N.; Oguchi, A.; Aoki, K.; Kubota, K.; Nakamura, Y.; Nomura, N.; Sako, Y.; Kikuchi, H.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix K1.

A:Reference number: A72450; MUID:99310339

A:Accession: B72719

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-110 <KAW>

A:Cross-references: DDBJ:AP000059; NID:g5103911; PIDN:BAA79246.1; PID:g5103930

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0291

C:Superfamily: bacterioferritin comigratory protein; alkyl hydroperoxidase c22

protein homology

B72719 Length: 110 May 30, 2002 09:26 Type: P Check: 4613 ..

1 MLVSGDPAPD IEIQLIDGST IRLSQLRGS VLYFYPKAF TPGCTREAIG
51 FNGLYEEFKK LGAEVIGVSM DPPGRNRRRA QNYGVRRFLA SDVEGEAFKS
101 FWGAQRPRTD

!!AA_SEQUENCE 1.0

PI:G70077 - hypothetical protein yxix - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: G70077

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Conerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.; Guiseppl, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Portwolik, S.; Prescott, A.M.; Presecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Taccioni, E.; Takegl, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.; Wambuit, R.; Wedler, E.; Wedler, H.; Welzenegger, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yaman, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033
A;Accession: G70077
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-153 <KUN>
A;Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB15950.1; PID:g2636460
A;Experimental source: strain 168
C;Genetics:
A;Gene: yxik
C;Superfamily: Bacillus subtilis hypothetical protein yxik
G70077 Length: 153 May 30, 2002 09:26 Type: P Check: 7683 ..
1 MEITSISIG NLDMIELKPD QTVMAELEED AESFYRWAG LAYDRIMIQV
51 ITTGSFIEDL SEYFEGHAYK VTKLAKREFH FQSILOEADR DIADFLFLA
101 SINDVFLIT DPQPKSYFS ESKLOCLTDS GERITWEYD AVDIWIGGE
151 SYK
1;AA_SEQUENCE 1.0
P1:T29895 - hypothetical protein F38A5.13 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Apr-2000
C;Accession: T29895
R;Fulton, B.; Stellyes, L.
submitted to the EMBL Data Library, September 1996
A;Description: The sequence of C. elegans cosmid F38A5.
A;Reference number: Z20705
A;Accession: T29895
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-589 <FUL>
A;Cross-references: EMBL:U70854; PIDN:AAB09157.1; GSPDB:GN00022; CESP:F38A5.13
A;Experimental source: strain Bristol N2; clone F38A5
C;Genetics:
A;Gene: CESP:F38A5.13
A;Map position: 4
A;Introns: 28/3; 121/1; 299/3; 345/3; 466/3

C;Superfamily: mouse Id-associated protein 1; dnaJ amino-terminal homology; myb DNA-binding repeat homology
T29895 Length: 589 May 30, 2002 09:26 Type: P Check: 5664 ..
1 MTTGNLOVAI YGFTARSRAF EPAGICYETR LIRDKLTLGH CTLPYKNSP
51 VAKPVEQQR KKSSEPAEL RKDLFDADNE KYEYLMKLD PNDCKNDHY
101 KVLGLSKLRW QATSDEIRRC YRQVKLKHHP DKKKHGIVM EKEEYFTCTP
151 KAYEQVMSD VKRQAFDSVD HKENDIIPNE KSIHNHFNYN ELAPVEQLNS
201 RWSNKKPVE LKSDATRED VENFYDFWPN FQSWREPSYL DEEDKERGED
251 RYERREMEKQ NKAERERRR EAKRIRKLIV DIAYAKDPRI IKFKKEQQAQ
301 KDKAKEDKOR AIREKQEAID REKREKEBAE AKQKEADRK AKEREERKK
351 ERDIKKAMS QQRKRLKILA DEAGHWTEPN RDKLTMEMERI ERICIGFTVD
401 QRELCEKVE SLSIASEIQT ALTDAELIKK EAAGAKYTIT EDKNKENEKQ
451 ADKETWTSEE IQLLVKASNT FPPGTVERWV QIADYINEMR KDSTGLPPKT
501 EKOVIKQCKA VQTMNVKLPS TTQNLGTPAL PDEDWSATE OKTLDEAIKK
551 HKSSDPERWE KISTEVGTRS KKACIRRFXY LVQMVKNKK
1;AA_SEQUENCE 1.0
P1:T19595 - hypothetical protein C31A11.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C;Accession: T19595
R;McMurray, A.
submitted to the EMBL Data Library, December 1996
A;Reference number: Z19149
A;Accession: T19595
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-705 <MIL>
A;Cross-references: EMBL:Z83218; PIDN:CAB05687.1; GSPDB:GN00023; CESP:C31A11.5
A;Experimental source: clone C31A11
C;Genetics:
A;Gene: CESP:C31A11.5
A;Map position: 5
A;Introns: 23/3; 86/1; 133/3; 287/1; 381/1; 474/2; 547/2; 579/1; 627/3; 656/2
C;Superfamily: Caenorhabditis elegans hypothetical protein F09B9.1
T19595 Length: 705 May 30, 2002 09:26 Type: P Check: 676 ..
1 MKIAPLFFI VCLFATVCE NMKEVFLAKA ADQKITGVSE QCANDTETWQ
51 KSLKMAVELS AECLIEQKCT KEELKTIEDN FYAVEQYDAW GKIPILGLFQ
101 LPILFDSYQ ECEIRISGRKY ATNYCYWVLM PGKNATCHMS DGLPTTFEER
151 GAVCMPYSCS EQDLPTVYNQ VSDQPFACA AFCSSYVKK TPAFWGFTSF
201 MAVMIGIALL ATVIDYLKDA LKKEDEKRED SRILQILITF SLWTNAELL
251 SVKEQKPGFI KCLDCIRFLS MLMVVTGHTF SYLTTPDQIE SILPFGGRFW
301 NHLVMNAFYS VDTFFLLSGL VVSYLEFKTK LKVSQISPI TWILFYVHRY
351 LRLTPPLMEF LGFFVYVGRY FQPGVASQL NQNGEVDTC QTYWKNLIT
401 INNLMSGDTQ CYGITWYLG A DTQLYLVAPI FLIGLYPSFA IGTALLTAAT
451 IGSVITVYIL FSTYDLPADF FGNGDATHFY DMIIYIKPMIR CPPYFGILV
501 GYLATYGRK KLRNLWALAV TGWIVAFSLG ALCIFSTYDY DNKVKWSIFS

551 RATYTNFSRL AMSFALSWI VANHMGWGP IDAFMSHPMW QPFGRLSYCA
601 YIVHYVVLVM YLMIGDASIH FYSSFOIFMY YAVPTTVLSY IFAFWSCLF
651 EIPFLKLEKM LIEILIGGAR DRNREDIEKQ KTLTKTKENE LMAVEETVOS
701 TNEKE

!!AA_SEQUENCE 1.0
P1:B81828 - hypothetical integral membrane protein NMA1989 [imported] -
Neisseria meningitidis (strain Z2491 serogroup A)
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: B81828
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morelli, G.; Basham, D.; Brown, D.; Chillingworth, T.; Davies, R.M.; Davis, P.; Devlin, K.; Feltwell, T.; Hamlin, N.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, M.A.; Rutherford, K.M.; Simmonds, M.; Skelton, J.; Whitehead, S.; Spratt, B.G.; Barrell, B.G. Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556
A:Accession: B81828
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-172 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85209.1; PID:g7380620; GSPDB:GN00124; NMA5P:NMA1989
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1989
C:Superfamily: Neisseria meningitidis hypothetical protein NMB1733

B81828 Length: 172 May 30, 2002 09:26 Type: P Check: 492 ..

1 MKDSLMDAFD FFMQGNFKKI ATEQGLDRKY FIFQYVCNL SAISSFCIMK
51 IIHTYTLIYI LIAVIGIING FLIRSKVLSI VKNKQKFLSD MFPLEMPFEF
101 VEKIPILAR FLNRGMKMLY LPSEFCLVAI CFSIYLNEE GGMPYFACY
151 WFGLCYPVST YVFLSRDYK HI

!!AA_SEQUENCE 1.0
P1:D81049 - hypothetical protein NMB1733 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: D81049
R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Ketchum, K.A.; Hood, D.W.; Peden, J.F.; Dodson, R.J.; Nelson, W.C.; Gwinn, M.L.; DeBoy, R.; Peterson, J.D.; Hickey, E.K.; Hitt, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Mason, T.; Ciecko, A.; Parksey, D.S.; Blair, E.; Clifton, H.; Clark, E.B.; Colton, M.D.; Utterback, T.R.; Khouri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizzi, M. Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venter, J.C.
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: D81049
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-167 <TEF>
A:Cross-references: GB:AE002523; GB:AE002098; NID:g7226980; PIDN:AAF42078.1; PID:g7226988; GSPDB:GN00119; TIGR:NMB1733
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1733
C:Superfamily: Neisseria meningitidis hypothetical protein NMB1733

D81049 Length: 167 May 30, 2002 09:26 Type: P Check: 3936 ..

1 MDAFDFFMÖG NEKRIATEÖG LDRKYFIFÖY IVCNLSAISS FCLMKIIRTY
51 TILYILIAVI GINGFLURS KVALSVKNKQ KFLSDMFPLE MPFFVEYKI
101 PILARFLNRG MKMLYLPSEF CLVAICFSIY ILNEEGGMP YFACYWFGLC
151 YPVSTYVFL SRDYKHI

!!AA_SEQUENCE 1.0
P1:A81923 - probable membrane protein NMA0784 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: A81923
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morelli, G.; Basham, D.; Brown, D.; Chillingworth, T.; Davies, R.M.; Davis, P.; Devlin, K.; Feltwell, T.; Hamlin, N.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, M.A.; Rutherford, K.M.; Simmonds, M.; Skelton, J.; Whitehead, S.; Spratt, B.G.; Barrell, B.G. Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556
A:Accession: A81923
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84067.1; PID:g7379505; GSPDB:GN00124; NMA5P:NMA0784
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0784
C:Superfamily: Neisseria meningitidis hypothetical protein NMB1733

A81923 Length: 173 May 30, 2002 09:26 Type: P Check: 1872 ..

1 MKDSLMDFFD FFAÖGKYKKI ATEQGLDRKY FIFHYVCNL SAISSFCIMK
51 VIHSYTLICL LIMIGIAGF FLIRNRVLSI VKNKQKFLSD MFPLEMPFEF
101 RYKIPILAR FLNRGMKMLY LÖNEFCLVMI CFSISILNE ERGMPYFIC
151 YWFGLFYPAN TYVFLSRDY KRI

!!AA_SEQUENCE 1.0
P1:S28025 - light harvesting complex II assembly factor pucc [imported] - Rhodobacter sphaeroides
C:Species: Rhodobacter sphaeroides
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 28-Jul-2000
C:Accession: S28025; T50701
R:Gibson, L.C.D.; McGlynn, P.; Chaudhri, M.; Hunter, C.N. Mol. Microbiol. 6, 3171-3186, 1992
A:Title: A putative anaerobic coproporphyrinogen III oxidase in Rhodobacter sphaeroides. II. Analysis of a region of the genome encoding hemF and the puc operon.
A:Reference number: S28023; MUID:93086425
A:Accession: S28025
A:Molecule type: DNA
A:Residues: 1-459 <GIB>
A:Cross-references: EMBL:X68796; NID:g46438; PIDN:CAA48701.1; PID:g46441
R:Choudhary, M.; Kaplan, S. Nucleic Acids Res. 28, 862-867, 2000
A:Title: DNA sequence analysis of the photosynthesis region of Rhodobacter sphaeroides 2.4.1.
A:Reference number: Z25222; MUID:20115911
A:Accession: T50701
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-459 <CHO>
A:Cross-references: EMBL:AF195122; PIDN:AAF24245.1

A:Experimental source: strain 2.4.1
C:Genetics:
A:Gene: pucc

S28025 Length: 459 May 30, 2002 09:26 Type: P Check: 7732 ..

1 MSRIAEHLVR IGRFLPFAD AASDQLPRK LLRLSLFOVA VGMAIVLVG
51 TLNRVMIVEL KVPASVVGIM ISLPLLFAPF RALIGFKSDT HVSALGWRRV
101 PWYRGTAL WGGFAIMPFA LIVLGGGGA EGQPFWLGS SALAFILWG
151 GGVHTIQTVG LALATDLAPR EDQPKVVGIM YVVLISMIF ASIGFGWLLD
201 PYDAQLIKV ISGVAVAVF LNMIALWME PRNRAFTVKP EKEPEFGDH
251 REFISRENAL HGLIVIGLGT LGFGMADVIL ERYGGEVLMS TVAETRLTA
301 TFAGGGLVGF WLASWVLGRG FDPRLMAFLG AAAGLPGFEA IMGATEMTNV
351 WVFLLGLTV GFGGGLFSHG TLTATMRLAP KEQVGLALGA WGAVOATAAG
401 VAAGAGVLR DILQAMPDLS GYGPAPYVA VFALPAGFLF LTMIVILPL
451 RSALAARRL

11AA_SEQUENCE 1.0

P1:A64690 - competence locus E - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: A64690

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sulton, G.G.;
Fleischmann, R.D.; Ketchum, K.A.; Klenk, H.P.; Gill, S.; Dougherty, B.A.;
Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus,
B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, K.;
Fitzgerald, L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne,
J.D.; Utterback, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman,
J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.;
Fraser, C.M.; Venter, J.C.

A:Title: The complete genome sequence of the gastric pathogen Helicobacter
pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: A64690

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-417 <TOM>

A:Cross-references: GB:AE000636; GB:AE000511; NID:g2314517; PIDN:AAD08401.1;
PID:g2314527; TIGR:HP1361

C:Genetics:

A:Start codon: GTG

A64690 Length: 417 May 30, 2002 09:26 Type: P Check: 8156 ..

1 MGVFLSILL AINLYLEYLN YQKLDSEKPT SLAQILLQY PKTKDQTYE
51 VLKLSKNMI FTTTIKEPLK NLQYRHAQFF GKIKPCSFLE SLKSCFFQTY
101 SFSLTRKQDF KSHWRHFIDS AHENALVGNL YRALFTGDSL NKDLDRANA
151 LGINHLLAIS GFHLGILSVS VFLLSLFYT PLQKRYFPYR NAFYDIGVLV
201 WVFLLGYLLL LDPLPSFFRA FLMGLLGFLA CFFGVRLLSF KLILACCIA
251 IALLPKLLFS VGFLLSVCGV WYIFLFLKHT QIEFKTSSFL MRSEQAISLS
301 ALVFLNMLLI VHAFFPMFSP YQLESIPGL IFIVEFPLSL FLHAVGLGSL
351 LDRLSLMPLT IPTISVPSPL WLGVLHLELT ILSARFEKVY LSMNVLSAGF
401 FLYCCYQYII MPSLIVG

11AA_SEQUENCE 1.0

P1:S74540 - hypothetical protein s110696 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S74540

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.;
Miyajima, N.; Hirose, M.; Sugita, M.; Sasamoto, S.; Kimura, T.; Hosouchi,
T.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo, S.;
Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium
Synechocystis sp. PCC6803. II. Sequence determination of the entire genome and
assignment of potential protein-coding regions.

A:Reference number: S74322; MUID:97061201

A:Accession: S74540

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-290 <KAN>

A:Cross-references: EMBL:D90899; GB:AB001339; NID:g1651650; PIDN:BAA16692.1;
PID:g1651764

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June
1996

S74540 Length: 290 May 30, 2002 09:26 Type: P Check: 1310 ..

1 MPKSTVNFNF NOQAIQTLTA PWRTGVFIYM LAAAWLPFLV PLSLAIDAN
51 LRSIVVMGIL FLIPVLVLI FWSHWCYTPL SLKAYGVYGL GWNRRQGVEL
101 LRGLGLFSF TEGLEFIQGL LGWAVLAPAG DRLMTIIMOG SLTGLGVALA
151 EELFFRGWLL KELEGGYGNK TSLASNAIIF AVLHFLKPLG EYVIRTLPOFP
201 ALVLGLSLG ITRRHGDR L GHSIGLHGM VMAYIIVNG QLVYITEKMP
251 AWVTGIDRNP LSGVVGIAGL CLLMLVNOG EKPMLEKRLGF

11AA_SEQUENCE 1.0

P1:S73732 - MG306 homolog A05_orf395 - Mycoplasma pneumoniae (strain ATCC 29342)

C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342

C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999

C:Accession: S73732

R:Himmelfreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae.

A:Reference number: S73327; MUID:97105885

A:Accession: S73732

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-395 <HIM>

A:Cross-references: EMBL:AE000039; GB:U00089; NID:g1674082; PIDN:AAB96054.1;
PID:g1674089

A:Note: the nucleotide sequence was submitted to the EMBL Data Library,
November 1996

C:Genetics:

A:Genetic code: SGC3

S73732 Length: 395 May 30, 2002 09:26 Type: P Check: 3483 ..

1 MSTPKSASFF TRKNILAFSE FIAFLVVSV LVTVEFLDIK TGDVKTIIINT
51 INRTNMPWIL LIVLGIIVTL AMNIIINMWV ARRCFHPAW WENVLFAQV
101 QFFQIVTPLS LGQDERLYW FIKKGMKQOT AVLLVSTGA FWNLAQALIT
151 WPSFEVLQN YALLENHNEG FVSYSWFSAG MIFDVVAIL FIFIAYSKR
201 HVLISGVNQ FRKWKRPYL TKQIYQRFI DKAENKLYG LEIKRLGLTI
251 FKLLANILIA VVGYSVEAV FAIVKENAT NNVIDQYSTA DIFNITINAI

301 TASNFIPIVPS GEGATQFVMT SELNAFKSAV GIESQVKGCV FLWRFLSYVI
351 PAIFSLCFI GWNVQVVEEF KHPKPVLPYV NLINHEWNN KKLHN
!!AA_SEQUENCE 1.0
P1:S46356 - pupR protein - Pseudomonas putida
C:Species: Pseudomonas putida
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 15-Jun-2001
C:Accession: S46356
R:Koster, M.; van Klompenburg, W.; Bitter, W.; Leong, J.; Weisbeek, P.
EMBO J. 13, 2805-2813, 1994
A:Title: Role for the outer membrane ferric siderophore receptor PupR in signal
transduction across the bacterial cell envelope.
A:Reference number: S46355; MUID:94298771
A:Accession: S46356
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-324 <KOS>
A:Cross-references: EMBL:X77918; NID:g525258; PIDN:CAA54871.1; PID:g525260
C:Superfamily: Pseudomonas putida regulatory protein pupR

S46356 Length: 324 May 30, 2002 09:26 Type: P Check: 5380 ..
1 MNGGATSIIP GEVAEQAMHW HLELOEPAVS AATLAACMSW ROAHPLHEHA
51 WQRTQVFAQR LREMRSPGQR PLAAHALRPQ QSRRTALKOL SLMAAGAGA
101 WYKDAALVQ DWRADYHSRI GEQRRLTLAD GTQVQLNTDS ALNVAFDQQA
151 RRLRLVRGEM LITRPALADS RPLWVDTEHG RLESTLAQFN VRLHGQHTQA
201 TVYQGSVALQ PALHAYPIL LGAGEQASFN QOGLLARQAV AAVAPAWSQG
251 MLVAQGQPLA AFIEDLARYR RGHLCADPAL AGLRVSGTEP LENTDKIITA
301 VAETLOLEVO HETRYWVTLK PRMA

!!AA_SEQUENCE 1.0
P1:A11701 - hypothetical protein homolog lin2155 [imported] - Listeria innocua
(strain Clp11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: A11701
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.;
Bloeker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve,
E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.;
Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.;
Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.;
Jackson, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.;
Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos,
B.; Perez-Diaz, J.C.; Rimmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes,
N.; Tlerraz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: A11701
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC97385.1; PID:g16414669; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin2155

A11701 Length: 391 May 30, 2002 09:26 Type: P Check: 3736 ..
1 MKATGIVEY NPFHNGQLH LNKARELTKS DVVIAVMSG FVQGEPAII
51 PKWERTKAL AAGVDVIEL PVSFATQHAT IFAESTIRLL DALHIDTLFF
101 GSEHGSEDF STAAKTIVEN EAAFNETIQL ALGDKRTSYA RAYTDAFTHL

151 FCKELLDVTK PNNILGFHYA LAIQKQNPST ALQTMPREHS GYHDAEASHD
201 YIASATAIRK LLAGNLEEA SRYLPDSSIE VLNNYRGPEL SFEDYWPILK
251 FRLIQASSDE LEGIRGVSEG IQNRMQLAOK KAHFSDFIE IMTKRYSNA
301 RLQRTALQIL LNAQNTPPAE PYIRVLGMSK TGQKYLSLHK KNISLPVYTT
351 VSKAEPSSLK EDLRATDIYT LINGLEDYQA GDFHTPPILT L

!!AA_SEQUENCE 1.0
P1:S28472 - rfbG protein - Vibrio cholerae
C:Species: Vibrio cholerae
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 08-Oct-1999
C:Accession: S28472
R:Manning, P.A.
submitted to the EMBL Data Library, May 1991
A:Reference number: S28467
A:Accession: S28472
A:Molecule type: DNA
A:Residues: 1-460 <MAN>
A:Cross-references: EMBL:X59554; NID:g48381; PIDN:CAA42138.1; PID:g48387
A:Experimental source: strain 017
C:Genetics:
A:Gene: rfbG

S28472 Length: 460 May 30, 2002 09:26 Type: P Check: 8871 ..
1 MTEKKLVVW VPLPPNSSWR GEGIAQTEN IVRNISPERK IEIVSSKHA
51 EMLVGLKSN PNISVLTIGF RKGSTKKTIG YVSLNEVERD SLMDLVIAKL
101 PIIPAIFRAV GMYVSQLEYL LSLYIYSHLQ RRGFFSSNCC RYWLPTPIIP
151 YTHLLGGEKF VSFWDPPVEE YNKEFPLTAE YFVKLLSKHF SNASAIITQS
201 RANKDYLETV MGIESSKINV IYNGSPDYSE FKKQGSNLSF SEVWSKSEFS
251 GASKKAFFEA LVNHQLNEFV LWRLLTKNKV SNRKIVLIST QNRPYKGFDO
301 LFLVLINECL RRDNYDFIFT CNVPTKLKER YPSLYERHIE VTRVDNYLHA
351 SLYLMSDIIV HPSNVEGGLG AYPQYEASSV GKPSLINTGR HVNEMAEGBF
401 DVDLLSSNFV NTKETVDKIE KLINSEFYMR QNIDAINRLK ISWKESASNY
451 ENVEFFGNENA

!!AA_SEQUENCE 1.0
P1:D82346 - rfbG protein VC0245 [imported] - Vibrio cholerae (strain N16961
serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82346
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.;
Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill,
S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.;
Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.; McDonald, L.;
Utterback, T.; Fleischmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.;
Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: D82346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-463 <HEIT>
A:Cross-references: GB:AE004113; GB:AE003852; NID:g9654648; PIDN:AAF93421.1;
GSPDB:GN00126; TIGR:VC0245
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:

A:Gene: VC0245
A:Map position: 1
D82346 Length: 463 May 30, 2002 09:26 Type: P Check: 1939 ..
1 MNVWTEKKKL VVWVPLPPNS SWRGEGLAQT IENIVRNISP ERKIEIVSS
51 KHAEMLVGLE KSNPNISVLT LGFRGKSTKK TIGYVSLNEV EKDSLMDLVI
101 AKLPIIPAIF KRVGMYSQ L EYLLSLYIS HLQRRGRFSS NNCRWMLPTP
151 IIPYTHLGG EKFSWFDPF VFEYNKEPPL TAEYFVKKL S KHFSNASAI I
201 TQSRANKDYL ETVMGIESSK INVIYNGSPD YSEFKKQSN LSFSEVWSKS
251 EFSGASKKAA FEALVNHLN FSVLWRLTFK NKVSNRKIVL ISTQNRPKG
301 PDQLFVLIN LCLRNDYDF IFTCNVPTKL KERYPSLYER IHEVTRVDNY
351 LHASLYLMSD IYLHPSNVEG GLGAYPOYEA SSVGKPSLIN TGRHVNEMAE
401 EGFVDVLSS NEVNTKETVD KIEKLINSEE YMRONIDAIN RLKISWKESA
451 SNYENVEFGN ENA
11AA_SEQUENCE 1.0
P1:E83776 - hypothetical protein BH1013 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: E83776
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirama, C.; Nakamura, Y.; Ogasawara, N.; Kuhara, S.; Horikoshi, K.
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis.
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E83776
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-155 <STO>
A:Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BAB04732.1; GSPDB:GN00137
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1013
C:Superfamily: Bacillus subtilis hypothetical protein yfjM
E83776 Length: 155 May 30, 2002 09:26 Type: P Check: 7616 ..
1 MNTLCIIPCG NKRIWDKQED IGPVAASEAY IGTILRLCAM YAEFFDHWV
51 ILSAKHGFLF PDDVVDGPYD VSFSTHSGV ITFDQLTEQV HEKKLDSVPH
101 VVLLTGKKYR PIVEACFPPEA TVECPPLSYS GIGYMQQALK RAIEQKQPLH
151 SAKRN
11AA_SEQUENCE 1.0
P1:S55948 - hypothetical protein YLR392c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein L8084.13
C:Species: Saccharomyces cerevisiae
C>Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 29-Oct-1999
C:Accession: S55948; C36445
R:Du, Z.
submitted to the EMBL Data Library, January 1995
A:Description: The sequence of S. cerevisiae cosmid 8084.
A:Reference number: S55944
A:Accession: S55948
A:Molecule type: DNA
A:Residues: 1-518 <DU2>
A:Cross-references: EMBL:U19729; NID:g625097; PIDN:AAB82352.1; PID:g625110; MIPS:YLR392c

A:Experimental source: strain S288C (AB972)
R:Ackerman, S.H.; Tzagoloff, A.
J. Biol. Chem. 265, 9952-9959, 1990
A:Title: ATP10, a yeast nuclear gene required for the assembly of the mitochondrial F-1-F-0 complex.
A:Reference number: A36445; MUID:90277691
A:Accession: C36445
A:Molecule type: DNA
A:Residues: 1-210 <ACK>
A:Cross-references: GB:J05463; NID:g1431794; PIDN:AAB05631.1; PID:g1480201
A:Experimental source: strain D273-10B/Al
C:Genetics:
A:Map position: 12R
S55948 Length: 518 May 30, 2002 09:26 Type: P Check: 3922 ..
1 MAPKISISLN PPNNGEYSS NDQMSGIVSL QLTKALSIRK ISVILKGFSE
51 TLTKIDQEYM FQONGMMMPG QDNKSFHTLM KFEQREPPD NYWVALDGSS
101 KPFKVPGSY NYSFQDFKP RKPECLKNHT AKTYAFVTRS NARLPPTFNS
151 HMQEFNKIDN LDLYEYSFGK VTYMVQVLE LGRSSSWFKP FHKLIREIET
201 FEFITPEPKDL IIEPDEDNE ELNAFSSNSR GNSWVTNNEF FNSNLSKVP S
251 KDVKVNVGVG YIKSDRNFSQ ANSILIENG D IRSRPVSSVT STROSTRLVN
301 GKVPFSTYK MGLPDGESNM RIEVSRDLK QIYRKDYLF R SGSQNFDPKYV
351 VMMEGNIASL SKMQITPLKL QLNLETTTY LSQGIANGNY SSLKLEIDL
401 NQLKSNKPLL DLNEIENFD GSMECELR L KDHPILRKLY FNEEDYRHRG
451 NRLYSRTCT IKRTSLQLL IEWGINGIRK QSEVVIDPVQ IFQVREHVE
501 AEALPRYVP PTYTEMAS
11AA_SEQUENCE 1.0
P1:F82557 - hypothetical protein XF2449 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82557
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing and Analysis, Sao Paulo, Brazil.
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: F82557
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-544 <SIM>
A:Cross-references: GB:AE004053; GB:AE003849; NID:g9107631; PIDN:AAF85248.1; GSPDB:GN00128; XFSC:XF2449
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvaranga, R.; Alves, L.M.C.; Araya, J.E.; Bala, G.S.; Baptista, C.S.; Barros, M.H.; Bonaccorsi, E.D.; Bordin, S.; Bove, J.M.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.; Colauto, N.B.; Colombo, C.; Costa, F.F.; Costa, M.C.R.; Costa-Neto, C.M.; Coutinho, L.L.; Cristofani, M.; Dias-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme, M.; Furlan, L.R.; Garnier, M.; Goldman, G.H.; Goldman, M.H.S.; Gomes, S.L.; Gruber, A.; Ho, P.L.; Hoheisel, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laligret, F.; Lambais, M.R.; Leite, L.C.C.; Lemos, E.G.M.; Lemos, M.V.F.; Lopes, S.A.; Lopes, C.R.; Machado, J.A.; Machado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.L.
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.;

Miyaki, C.Y.; Montei-ro-Vitorello, C.B.; Moon, D.H.; Nagai, M.A.; Nascimento, A.L.T.O.; Netto, L.E.S.; Nhani Jr., A.; Nobrega, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Paris, A.; Peixoto, B.R.; Pereira, G.A.G.; Pereira Jr., H.A.; Pesquero, J.B.; Quaggio, R.B.; Roberto, P.G.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki, H.E.
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira, J.F.; Silvestri, M.L.Z.; Siqueira, W.J.; de Souza, A.A.; de Souza, A.P.; Terenzi, M.F.; Truffi, D.; Tsai, S.M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zago, M.A.; Zatz, M.; Meidanis, J.; Setubal, J.C.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2449

F82557 Length: 544 May 30, 2002 09:26 Type: P Check: 144 ..

1 MICRRYLVLD SMNLRFSHW FVELLSVAL LLGCKQDSP SAKROIIVE
51 AKLPPRLKP DLPLFLDIE RRTQFFWDT TNEVNGLTDP RYPSRPFASI
101 ASVGALATAY PIGIENGWIS RNOAIDRTLT TLKFLRDAPM GPQRTGCAGY
151 KGFYHFLDM QHGRYDSWV ELSSVDTALL MGVLFVESY YDGEDVREKE
201 IRQIADKLIR RVDWLYLQQR KPLISMGWYP ERGFIEHDM GYNEAMVYL
251 LALGSPTHPL EPVSWDEWTR TYNKDWGVFQ GQEYLAFGPL FGHQYTHWV
301 DFRDIQDQYM RERGIDYFLN SRAVLNHRD YAINPMKWK DYGENVWGLT
351 ASDGPQNTQ EYRGEQRQFT HYRSRGAGLT EFPDDGTAP TATVASIVEA
401 PEVVIPTATLE MHKRYGDFLY SSYGFLDAFN PSFDYNVPLK TGRLIPGRW
451 VASDYIGIDQ GILSMIANY RNDFWNVWK KKKYVRTGLE RAGFIGWILQ
501 EDEVLRGPQR DERAAYRSL GIAESHAPPA QTORVPSSLG KKPE

IIAA_SEQUENCE 1.0

P1:F82112 - hypothetical protein VC2147 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: F82112

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleischmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: F82112

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-118 <HEI>

A:Cross-references: GB:AE004287; GB:AE003852; NID:g9656689; PIDN:AAF95292.1; GSPDB:GN00126; TIGR:VC2147

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2147

A:Map position: 1

F82112 Length: 118 May 30, 2002 09:26 Type: P Check: 2237 ..

1 MKNTTLPSE QTOQEAALIA KATQPGQTK EQTKLITQGI EKGIALYKKQ
51 OKEKHRADK LRKALKAKQ SSTEIHEAD DYAAELGDDS HTNQAKLAWV

101 LIALSWLGFI SYWLWONS

IIAA_SEQUENCE 1.0

P1:C84898 - hypothetical protein At2g46060 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: C84898

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.
Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: C84898

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-787 <STO>

A:Cross-references: GB:AE002093; NID:g3702340; PIDN:AAC62897.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g46060

A:Map position: 2

C84898 Length: 787 May 30, 2002 09:26 Type: P Check: 3197 ..

1 MAELFLLR TLVFSICAL ICNCLQDAAD NNGNEFTVSS FRYPESEVRP
51 YDTRYIRVDL PPWFSSLNVA IESVDVITAK SISKISKSL PVICFRDGP
101 PLPDASTNAL QGLELGRFFN GSFERAQDSE IAQOCYPMOK NITRLTNEQ
151 CLEVLILISY NLDTAQIVRS SAFSPSANIS VEGCKTATMW GPFCNQTIYP
201 LSCSRFDNOT ASVISCADSF PSSCLTGAET KYTALDVDCI AEQLVIMASN
251 VKVDSNESYL MCFARFEAF SETLHDYAAD IHKVPPIVWK PKAGRWTYVI
301 SLSGRENREFA QCTNSSSRVC FSINVKVLCG PVKAGPNCG QOIYILQAVM
351 RRGWLTPFOS YFPPVNDASL SGSSTNFPLE PIVSNFSSIP ELDSTWTYF
401 LMNIPQGGSG GHIHFRLLSD STIQEYVYLR FGLPTIDDR DYIYVNRISA
451 SRSMFESLYN SSKEMVDFYI LYAREGTWSF GLRQLIDSNT PAASRGSPPL
501 VSLSLERCPR GCSSYGQCRY AFDANGLTSY REFIYLESK ICKHNGNCS
551 CDRTHGGFDC SIEIVSHQEH IVQSIALIAS NAAALLPAYW ALRQREYPEW
601 VLETTSSGISS ALYHACDVGT WCVLSYNVLQ FMDFWLSFMA VVGTFVYIST
651 AGEAVKRTIH TVVAILTALL ALTOATRASN IIVLAIGSL GLLIGFLVEF
701 VTKYRSYCGS AGFSLNMLDR PRAVKEWFSN LIKTLKKRRR WGFVAAGIYA
751 FTMAAISFKI AIVNHENLAH NGADNYELTR QDSLRSN

IIAA_SEQUENCE 1.0

P1:C84582 - hypothetical protein At2g19880 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: C84582

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.

Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: C84582
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-550 <STO>
A:Cross-references: GB:AE002093; NID:g3687230; PIDN:AAC62128.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g19880
A:Map position: 2

C84582 Length: 550 May 30, 2002 09:26 Type: P Check: 6852 ..

1 MSTLDSIDAI LFSLSRAFTS PFAVFVQIQG CTICLLALG WLLAEYVRNR
51 EVKRIKNSIK AGNSLAFLYQ DINELEHSRQ VKLPRVSVM PLKGFEHNL
101 HNWRSQITSL YGCPLEFLFV VESTEDPAVH AVSRLLSMYQ DHEAKVVVA
151 GLSTCSQKI HNQLIGVERK HKDTKYVLF L DDDVRLHPGT IGALTTEMEK
201 NPEVPCSMGF ATGRTFFLW GCGMMHADD FRQDRYGVVS GLRDGYSDD
251 MTLASLAGAH KRLITSPVA VEPHPLASDL SFGRYWNYLR KQTFVLESYI
301 SKVNMINKA LFAVHCYLSW GFVAPYVMAI IHITSALRIY IKGYHOLEDT
351 TSASGGLSVF LMLHIDRRMR FYVGMLLVIT LAICTFIELL SMWNLTREV
401 QLCNMLSPEA PRLSLATYNN GLPFIIGHI SPMYRGFNKP RTGVFWNVED
451 YPIPNIYPHT IFQNIKSAL NKHFNVETPL LLQTVSSSIW LSTSLFDRGI
501 LIKPSFTSSA ASWPETMYK GLNGKTSLYL FARAIONIH SNIAWSARWC

!!AA_SEQUENCE 1.0
P1:AD0772 - hypothetical protein STY2348 [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18)
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AD0772
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; Mungall, K.L.; Bentley, S.D.; Holden, M.T.G.; Sebahia, M.; Baker, S.; Basham, D.; Brooks, K.; Chillingworth, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J.; Feltwell, T.; Hamlin, N.; Haque, A.; Hien, T.T.; Holroyd, S.; Jagels, K.; Krogh, A.; Larsen, T.S.; Leather, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.G.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.
A:Reference number: AB0502; PMID:11677608
A:Accession: AD0772
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-703 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02498.1; PID:g16503362; GSPDB:GN00176
C:Genetics:
A:Gene: STY2348

AD0772 Length: 703 May 30, 2002 09:26 Type: P Check: 9075 ..

1 MEQDDRLLNA MFEWCNHNK LNDGQREWHI ADIPGLREE RYDELDERYN
51 QALTESFTSR EAEKRYFFAW NOMDNPFYDM DTLVEAGPOG LALIKNMQRA
101 RPRSTHAWLA EAQYWNHRAW LYRSYWARE TTRAMWICAA ACNERMVIAA
151 LNAIDCEPRQ WMAAALTSTN SKVFGQPDWL VEFLEGADVA GQPLMEDIAE

201 YHRSPQEV D ALMAHSGLSF ADAVCPNLP R PSVLPECND D AGQKYWLAVC
251 LAIFPTAFYV LDEYIPFRMP RMGGSHEIR EFLESSVCDH ISAAREHLE
301 LLIMWDDHRD LRIKEVDSPA EQERITAKAE EISLRAHQE SRHNTLKLWR
351 VCYSDDLDDND ALWRTLQRSI VEKVFNNYF FDDTIKFLAR DEPDWMMYN
401 FLCQNAQOTE FAVPKIRRGY FOYAGLLGFE KDEAQGLAWL DSVADIQYNH
451 NMRAAIKNFD WFGLEPHFVP LAELGAQRNI PALNLLGLE HNNKENNGLL
501 PYDPAIALGY FQRAAEILHR RLALRESTPY KLIDNGGYTD YENDLQNIHF
551 SIGICNQRLS KQELDTEKRS AYEKELLDNL WLAHQFGHKE AMGLFLNIF
601 EVKDITLAKH HLELVQOEAN KGTJHAMVTL SRLHGKHDR TLFNMKLSAR
651 WAHFAFTLYP DNEIYMDCLD HLHFDSEFWKR FFEAWYTVRI PSELPGQVN
701 SMV

!!AA_SEQUENCE 1.0
P1:T20228 - hypothetical protein C54G10.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T20228
R:Matthews, L.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19240
A:Accession: T20228
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-660 <WIL>
A:Cross-references: EMBL:Z75532; PIDN:CAA99810.1; GSPDB:GN00023; CESP:C54G10.3
A:Experimental source: clone C54G10
C:Genetics:
A:Gene: CESP:C54G10.3
A:Map position: 5
A:Introns: 86/3; 133/2; 305/3; 336/2; 442/1; 628/3
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

T20228 Length: 660 May 30, 2002 09:26 Type: P Check: 7424 ..

1 MARLADLEF FRRFRWLKI LFPCHRRSIT AWLAGALLIA TALDOVMTYL
51 VGIDPSLEYV ALGNKADDTF KILCARGAAI ILGKAFTLAV FKYLTNMLAI
101 KSRQVCNLTM HRLYFKRQAF FKLISSGDM L DNPQRLTQD IEKATRILSN
151 DLLAPIATAP FIGIYYTWT L YESSGWIGPA AIITYFCIQ T IINKMILSPI
201 VQKVSQEK M EGDFFQRHME VRSNVEAIAF YRAGVLENIM TNQKLKLI E
251 TQKSLTEWRM VLNSITNVFD YFGGILSYLI IGVPEFITH L YDDVSPAELN
301 GIVSRNAFFY LYLITYSESTV LKLTGDFGEL AGVTHRMEL HEELNRLHSD
351 CLETRPPST VPSSVVIVAS DEDDKSASRH MQETIHKQMS LERDEQEEEE
401 AQYLGGKTG QEDDWPDDGV AITVDSATLS PRNDHSHLIV QLSLQIIQG
451 QTLITGDSG CGKSSLRMF AGIHWCSSGK MDCHWRRLTS NLFELAQKPY
501 FPSGNTLRQ QIVVPKALQ VDKDVARITQ ILEWVKMEHL VERCGGLDTP
551 VEWDMKTL S PGLQRLSLA RVEYTKPRIV FLDESTSAIG FELEMAIYRK
601 LQEKITFVS IGHRSYSLQF HDMLRVKGR SGEMSLHDID TASIASRTAS
651 FLGADTVLSM

1!AA_SEQUENCE 1.0
P1:H87334 - conserved hypothetical protein CC0691 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87334
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.; Alley, M.; Ohta, N.; Maddock, J.R.; Potocka, I.; Nelson, W.C.; Newton, A.; Stephens, C.; Phadke, N.D.; Ely, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonay, J.F.; Smit, J.; Craven, M.; Khouf, H.; Shetty, J.; Berry, K.; Uterback, T.; Tran, K.; Wolf, A.; Vamathevan, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87334
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-208 <STO>
A:Cross-references: GB:AE005673; NID:g13421912; PIDN:AAK22676.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0691

H87334 Length: 208 May 30, 2002 09:26 Type: P Check: 7467 ..

1 MPPIVDNGP MLNLTADILQ GRVRLPEVY VDHDELKAA IDCDPASMEI
51 MSVNGCGEGF EDWFGALGE TDRGERIGFA IRLVDEKRV GTSSYLNIIR
101 LHGLEIGAT FLNPEARSGP VNPESKRLML GHAFDKAGAI RVELVTDVNR
151 ARSQAAIQKL GATKEGVLRN HKVTWGHVR DTAVFSTIDY DWPAIRERLE
201 FRLSETFA

1!AA_SEQUENCE 1.0
P1:E86254 - hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E86254
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltl, R.; Marziani, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pal, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Uterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: E86254
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <STO>
A:Cross-references: GB:AE005172; NID:g3157932; PIDN:AAC17615.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

E86254 Length: 783 May 30, 2002 09:26 Type: P Check: 9544 ..

1 MATLGDIGVA AAINILTAII FLAFAILRI QPNDRYVFP KWLKGIRSS

51 PLHSGALVSK FVNVNLGSLY RLNMMPAAL KMPEELIDH AGLDSAVYLR
101 IYLGKIFV PIALAWSIL VPVNWTSGL QLAKLNRVTS SDIDKLISIN
151 IENGSDSLYF GRFWTHLVMA YAFTEWTCYV LMKEYEKVAA MRLAFLQNEQ
201 RRPDQFTNLG LSQLLSQVLV RNVPADPDES ISDSVEHFEL VNHDPHYLTH
251 QVYVNNANDLA ALVEOKKSTQ NWLDYQYLYK TRNQEHKPRI KTGFLGLMGK
301 KYDAIDHYIA EIEKLENOER KVKKKDDTSV MPAAFVSFKT RWGAAVSAQT
351 QQSSDPTEWL TEWAPEAREV FWSNLAIPYV SLTVRRLIMH IAFFLTFFEF
401 MIPIAFVQSL ASIEGIEKNA PELKSIEND LEKSVIQGFL PGIVLKPLFI
451 FLPSILMWMS KEEGFVSLSS LERRAERFY IFNLINVLG SVITGSAREQ
501 LDSEFLKQSAK EIPKTGVVAI PIKATFFITY IMVDGWAGIA GEILRLKPLI
551 PFHIKNSLIV KTEKDREEM NPGQINVHAT EPRIQLYFLL GLVYAPVTPV
601 LLPETIIFFA LAYLVRHQI INVYNQEYES AARFWDVHG RIISALIIAQ
651 ILLMGLSTK GAAQSTPFL FLPIITFFFH RYCKGRYEPA FLRHPLKEAM
701 VKDTLERARE PNFNLKPYLQ KAYIHPVFKD NDYEDSRFDE ISGYCIEDSD
751 EECVTVPTRK QSRINTPAVS HASRGSSRSP PSK

1!AA_SEQUENCE 1.0
P1:C96782 - unknown protein F22H5.13 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96782
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltl, R.; Marziani, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pal, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Uterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: C96782
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-654 <STO>
A:Cross-references: GB:AE005173; NID:g10092280; PIDN:AAG12693.1; GSPDB:GN00141
C:Genetics:
A:Gene: F22H5.13
A:Map position: 1

C96782 Length: 654 May 30, 2002 09:26 Type: P Check: 7509 ..

1 MSTTSSVRVR LAFVALLSAT TFCYIHKYRR LKHLKNLSLN PSSTLKASRG
51 KIFISQGTGT AKALAQRLEH LCASNDIAFD IVDPHSYEPE DLPKETLVLF
101 IASTWDGGRP PKNGEFLVNW LGESAEDFRV GSLLSDCKF AVFGVGSRAY
151 GESYNAVAKL LSSRMIGLG LEMIPVGEQ VDDGELDRAV QDWCDGIVRV

201 LKGSAGQETN GVSQOIGAVE DDLEYDSTD EEDEDNDADG GIVDLEDIAG
251 KAPSKRNGVV KVTKVDGKKE MVTPIRASL TKQHDFYLLQ GYKITGSHSG
301 VKICRWTKSQ LRGRCGCYKH SFYGIESHRC METTPSLACA NKCVCWRHH
351 TNPVGKSMQW KMDEPSIVK GADLHKHNM KQMGVPGVT PEKLOEGLNP
401 RHCALSLVGE PIMYPEINAL VDELHGRRIS TELVTNAQFP EKILMMKPIT
451 QLYVSVDAAAT KESLKAIDRP LEADFWERFI DSLKALQEKQ QRTVYRLTLV
501 KGNWTEELDA YFNLFISIGK DFIEIKGVTY CGSSATSKLT MENVPWHTDV
551 KARSEALSLK SNGEYEVACE HAHSCCVLLG RTEKFKVDGK WFTWIDYEKF
601 HDLVASGEFP TSTDYMAQTP SMAVYGAQEG GFDPQLRYK KERHHPKPKQ
651 AVLA

IIAA_SEQUENCE 1.0
PI:A84888 - hypothetical protein At2g45230 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: A84888
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.;
Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.;
Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.;
Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.;
Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.;
Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg,
S.L.; Fraser, C.M.; Venter, J.C.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: A84888
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1374 <STO>
A:Cross-references: GB:AE002093; NID:g2583130; PIDN:AAB82639.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g45230
A:Map position: 2

A84888 Length: 1374 May 30, 2002 09:26 Type: P Check: 8740 ..

1 MRILSWNCOG VGNTPTVRHL REIRGLYPE VIFLCETKKR RNTLENVVGH
51 LGFFDLHTVE PIGKSGGLAL MWKDSVQIKV LQSDKRLIDA LLIMQDKEFY
101 LPTCYGEPVQ AERGLWMLERL TRIGLSRSGP WMLTGDENEL VDPSEKIGCP
151 ARKSSCLEF RQMLNSGLW EVNHSGYQFS WYGNRNDELV QCRLDRTVAN
201 QAMMELFPA KATYLOKICS DHSPLINLV GDMWRKWAGF KYDKRWQRE
251 GFKDLLCNFW SQOSTKTNAL MMEXIASCR EISKWKRVSK PSSAVRIQEL
301 QFKLDAATKO IPFDRELAR LKKELSQEYN NEEQFWQEK RIMMMRNGDR
351 NTKYFHAAATK NRRAGNRIOK LIDEGREWTD SDEDLGRVAE AYFKKLFASE
401 DVGTYVEELE NLTPLYSDQM NNNLLAPITK EEVQRAFESI NPHKCPGPDG
451 MNGFLYQGFW ETMGDQITEM VQAFRRSGSI EEGMKNKNIC LIPKILKAEK
501 MTDPRISLC NVYKVIKGL MANRLKILP SLISETQAAF VKGRLISDNI
551 LIAHELHLAL SSNNKCEEF IAKTDLISKA YDVRWMPFLE KAMRGLGFAD
601 HMIRLIMECV KSVRYVLLIN GTPHGEIIPS RGLRQGDPLS PYLEVICTEM

651 LVKMLQSAEQ KNQINGLKVA RGAPDISHL FADDSMEYCK VNDEALGOIT
701 RIEEYSLAS GQRVNLYKSS IYFGKHISEE RRCLVKRLIG IEREAGEGVY
751 LGLPESRQGS KVATISYLLKD RLCKRYLGWQ SNPLSPGCKE ILLKAVAMAL
801 PTYTMSCFKI PKTICQIES VMAEFWKNK KEGRLHWKA WCHLSRPKAV
851 GGLGFKIEA FNIALGKQL WRMTTEKDSL MAKVEKSRYE SKSDPLNAPL
901 GSRPSFAWKS IYEAQVLIKQ GIRAVIGNGE TINWTDPMI GAKPAKAAQA
951 VKRSHLYSQY AANSIHVKD LLLPDGRDWN WNLVSLLEPD NTQENILALR
1001 PGKETFRDRF TWEYSRSGHY SVKSGYWMVT EIIQNRNNPQ EYLQPSLDPI
1051 FQQIMKLDVP PKIHHEFLMRC VNNCLSVASN LAYRHLAREK SCVRCPSHGE
1101 TVNHLEFKCP FARLTWALISP LPAPPGGEWA ESLFRNMHHV LSVHKSQPEE
1151 SDHHALIPWI LWRLMKNRND LVFKGREFTA PQVILKATED MDAMNNRKEP
1201 QPQVTSSTRD RCVKMQPPSH GWVKCNTDGA WSKDLGNGCV GWVLRNHTGR
1251 LWMGLBALP SQQSVLETEV EALRNAVLST SRFNRYRVIF ESDSQYLVSL
1301 IONEMDIPSL APRIDIRNL LRHFEEVKFO FTRECGNNVA DRTARESLSL
1351 MNYDPKMTSI TPDWIKNLVD LETV

IIAA_SEQUENCE 1.0
PI:A84448 - hypothetical protein At2g03400 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: A84448
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.;
Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.;
Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.;
Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.;
Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.;
Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg,
S.L.; Fraser, C.M.; Venter, J.C.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: A84448
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-131 <STO>
A:Cross-references: GB:AE002093; NID:g4335759; PIDN:AAD17436.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g03400
A:Map position: 2

A84448 Length: 131 May 30, 2002 09:26 Type: P Check: 6617 ..

1 MVQSQSLSTL TICGSYKVS LLRNRLNSVK ASSLIGDRCV SCQFLRKSPS
51 FRSHWKSLLQ RNLRLVEARW PFQGGEGEGL DPSSERSESA NEDILIFFEQ
101 LDLATRYQVK TKCRVFISFK MVANGFDVLQ Y

IIAA_SEQUENCE 1.0
PI:H83554 - hypothetical protein PA0736 [imported] - Pseudomonas aeruginosa
(strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83554
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey,
M.J.; Brinkman, F.S.L.; Huftnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber,
R.L.; Goltzy, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.;
Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.;

Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.
A:Reference number: A82950; MUID:20437337
A:Accession: H83554
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <STO>
A:Cross-references: GB:AE004508; GB:AE004091; NID:g9946611; PIDN:AAG04125.1; GSPDB:GN00131; PASP:PA0736
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0736

H83554 length: 358 May 30, 2002 09:26 Type: P Check: 7512 ..

1 MOPIDRAQAA QRANDILVFQ RELQRLDQEQ AFRLEPAQAR QLADYHLL

51 DSYRDRFDID HDLRSQQLSL GMRVASFEFGA LAFASVELL FYRFWGLFPT

101 VAOVAAILIGS AFAAFPAITLW VOAKDASGYF SKLAAMVAFV CFVLDTMIG

151 QIFNVTPSDL ALVPWALYAL LLAYLCNARL LLAAILCYM GFIAARVGTW

201 GGGYWLVSVE RPFNFPPAAA LIFAVPLCFD QRNFSGFAYI YRVFALLGLF

251 LPMVLNANWG SGSYLALPSA LIEGLYQVAG FVAALVITWL GARRNMADVS

301 NTGITFEVIF LTKFFDWMW EAMPKYLFFL VLGLSALLIL LVLRRLRTP

351 GIARTDA

!!AA_SEQUENCE 1.0
P1:F83173 - outer membrane protein OprC PA3790 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83173
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber, R.L.; Goltzy, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.; Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.
A:Reference number: A82950; MUID:20437337
A:Accession: F83173
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-723 <STO>
A:Cross-references: GB:AE004797; GB:AE004091; NID:g9949950; PIDN:AAG07177.1; GSPDB:GN00131; PASP:PA3790
A:Experimental source: strain PA01
C:Genetics:
A:Gene: oprC; PA3790

F83173 length: 723 May 30, 2002 09:26 Type: P Check: 9154 ..

1 MEKRMSTQOR AAGNACPTAA FSDPARLAQ RRRWAGAFPA ICGLALSPSA

51 LLAEHSQHQ DHAVALPSV VTGVAQSSPL TIVTNPKRPR QPVASDGAD

101 YLKTIPGFAV IRNGSGNDP VLKGMFGSRL NILTNGMML GACPNRMDAP

151 TSYISPETYD KLTVIKGPQT VLMWPGASAG TILFEREPR FGELGSRYNA

201 SLAGSNGRF DKVLDAAGN RLGYLRFNGN HAQSDYEDG AGNTVPSRWK

251 KWNQDVAVGW TPDEDTLIEL TAGKGDGEAR YAGRGDGSQ FKRESLGLRF

301 VKSNVSDVLE KVEAOYYINY ADHIMDNFRL RPPDPSSMMP MEMASQVDRR

351 TLGGRHAATW RWDFFKLVTG VDAKRNEHRA RSKYDMMTD YTTDADQFPW

401 SKDAVFHNYG AFGELTWFAA ERDRLIGLRL LDRASVKDYR QTLKSGHMGH

451 AMANPTANDT RADTLPSGFV RYEHDLADSP TLLYAGLGA ERFPDYWELE

501 SPKRPNGSV NAEDEKIKPEK TTQULDFGLQY NGDKLQAWAS GYGVVQDFI

551 LFSYREGMMG SSTQATNVDA RINGGELGAS YQLTGNMKTQ ASLAYAWGKN

601 SSDDRALPQI PPLEARFGLT YEEGDSWAGS LMRVAPQNR IARDQGNVVG

651 KDEKKSAGFG VESLNGAYRV TRNVKLSAGV DNLFDKYTE HLNKAGDAGF

701 GFSANETVPE PGRTFTTKVD FSF

!!AA_SEQUENCE 1.0
P1:F83610 - probable fatty acid desaturase PA0286 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83610
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber, R.L.; Goltzy, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.; Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.
A:Reference number: A82950; MUID:20437337
A:Accession: F83610
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-392 <STO>
A:Cross-references: GB:AE004466; GB:AE004091; NID:g9946120; PIDN:AAG03675.1; GSPDB:GN00131; PASP:PA0286
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0286

F83610 length: 392 May 30, 2002 09:26 Type: P Check: 1486 ..

1 MWFNGFLDLS PMQOLVAVTL LTHITIVSVT VYLHRYSAHR ALDLHPALQH

51 FFRFWLWLT GMNTREWTAI HRKHAKCET VEDPHSPVHK GLFTVLRAGA

101 ELYKAEAKNO DTLRIYKNC PDDMIERNLY SRPPIGVVL MALIDLALFG

151 ALGLTWAVQ MWIPIFWAAG VVNGLGHAVG YRNEECRDAA TNLVPWGILI

201 GGEBELNNHH TYPNSAKLSV KKWEFDLGA WIKLFSFLRL ARVARVAPIA

251 HRVEGKHSID MDTAMAILNN RQIMAQYRK LVIAPLVKQF VAKADESVRH

301 LFRRAKRLLS RETSLQDRH HVRIDSMLAH SQALKVIYER RLALQOIWAK

351 TSANGHDMLA AIKDWVHEAE ASGIQSLKEF AAQKTYSLR PA

!!AA_SEQUENCE 1.0
P1:S73657 - MG288 homolog F04_orf260v - *Mycoplasma pneumoniae* (strain ATCC 29342)
C:Species: *Mycoplasma pneumoniae*
A:Variety: ATCC 29342
C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
C:Accession: S73657
R:Himmelfeich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R. Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma*

pneumoniae.
A:Reference number: S73327; MUID:97105885
A:Accession: S73657
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-260 <HIM>
A:Cross-references: EMBL:AE000031; GB:U00089; NID:g1674002; PIDN:AAB95979.1;
PID:g1674006
A>Note: the nucleotide sequence was submitted to the EMBL Data Library,
November 1996
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG
S73657 Length: 260 May 30, 2002 09:26 Type: P Check: 3241 ..
1 MQAFKEYWQK QKDVTDKQ LLEALKLSFA KEQNKTFAPL IKNFQDGTISN
51 YYPNDQEDQS EAKTAFGTQ GIAFPQSGLK GIFMSEWLK QLGKAKINL
101 DIKSLKVTDS KISPTIKMK DIGIKRNQDK PYNFRFEIDI EYQGNKLSW
151 LEAIIAKFSG IPGEWKGKLN LKFTVDGDS WEIVQKPDYP GSLFQFDDOK
201 QQLFLKLIHW EKITVQEPF MELIKSONLH NLELRTESTK PPVVDLASYL
251 HYQLKLNQO
11AA_SEQUENCE 1.0
P1:E69365 - hypothetical protein AF0925 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: E69365
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.; Peterson, J.D.; Richardson, D.L.; Kerlavage, A.R.; Graham, D.E.; Kyrpides, N.C.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.; Peterson, S.; Reich, C.I.; McNeil, L.K.; Badger, J.H.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uitterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.; D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.; Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Archaeoglobus fulgidus.
A:Reference number: A69250; MUID:98049343
A:Accession: E69365
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-315 <KLE>
A:Cross-references: GB:AE001040; GB:AE000782; NID:g2689363; PIDN:AAB90325.1; PID:g2649684; TIGR:AF0925
E69365 Length: 315 May 30, 2002 09:26 Type: P Check: 7306 ..
1 MLESWATSSL ILFVALIVV SLSLLYCYL RRRMMWKISG IFSVASPAVY
51 ILLFWSEFN LLSNLFILL VELAVFIIL PIKVFPELK FENEKYGEL
101 DGIPVIIGYE KGKKVYNIFY TPLKRKIIFYT KSLKDVLSGE ELKAVIYHES
151 GHSKNKMMMI TRSTAMFEV LIAAVLTIL FLEMKGKFP NLKVSLEFILL
201 GALLIIVATF FWFVSWINEH EADLFAVAKS GYENFSKALF KTYFYNVLGD
251 YAEFVGKIDL KNFNSGDVTP FEILKILKQ SIYLLFPRNI LNQIPQTHP
301 PLRYRILLAH QTLKC
11AA_SEQUENCE 1.0
P1:T01809 - hypothetical protein A_TM021B04.3 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 22-Oct-1999
C:Accession: T01809
R:Dante, M.; Wamsley, P.; Gibson, A.
submitted to the EMBL Data Library, June 1997
A:Description: The sequence of A. thaliana TM021B04.
A:Reference number: Z14440
A:Accession: T01809
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-938 <DAN>
A:Cross-references: EMBL:AF007271; NID:g2191181; PID:g2191183; GSPDB:GN00063; ATSP:A_TM021B04.3
C:Genetics:
A:Gene: ATSP:A_TM021B04.3
A:Map position: 5
A:Introns: 125/1; 559/3; 664/1; 798/3
T01809 Length: 938 May 30, 2002 09:26 Type: P Check: 5621 ..
1 MKRENNLVLS LLEFVIVFLM QVGEAQNRIY NVNNGIVNDI GTAYSNMTLL
51 CINMSLSDFY SSHPETQTRL VTTVDSKND VVTAAGTR IPISLNHAL
101 RYVCVINHQF ICTLLIYIA FSSKALDLIT NKEVKAILGP WTSMAQFMI
151 EMGQKSQVPI VYYSATSPSL ASIRSQYFFR ATYDSSQVH AIKEIKLFG
201 WREVAAPVVD DTFEGCIMPR LTDVLQEIIV RPYRTVISP NATDEISVE
251 LRRMNTLPTR VEVVHLVELL ASRFEAKATE IGLMKQYVW ILTNTITDVL
301 SINNETEIEI MQGVLGKTY VPSKLELENF RSRWTKRFP I SDLVNYGLWA
351 YDATTALALA IEEAGTSNLT FVKMDAKRVN SELQIGVSO YGPKILQTL
401 RVRFOGLAGD FQFINGELQP SVEEIVNVNG QGGRTIGFWM KEYGLFKNVD
451 QRPASKTTFS SMQDRLRPII WPDGTTSPK GWEIPTNGKR LQIGVPVNT
501 FQGFVKATRD PITNSTIFSG FSIDYFEAVI QAIPIYDISYD FIPEQDGYD
551 ALVYQVYLK YDAVVADTTI SSNRSMYVDF SLPTPSGVG LVVPVKDSVR
601 RSSSTIFLML TLALMLISL SFFIIGLVW VLEHRVNPDF DGPQYQLST
651 IFWFSESIMV FAPRERVLSE WARVVIIWY FLVLVLTQSY TASLASLTLT
701 QHLHPTVTNI NSLLAKESV GYQSSFILGR LRDSGSEAS LVSXGSEHC
751 DALLSKGQAE GGVSAVLMEV PYRIFLGQY CNKYMVQTP FKVDGLGFVF
801 PIGSPVLADI SRAILKVEES NKANQLENAW FKPIDESCPD PLTNPDNPS
851 VSEFQLGFD FVWLFVLA I VCTMALLKFV YQFLKENPNQ RNLRLWMEKF
901 NEPDQKSYIK DVTKCQCSSG QGMPKNGQEG ANAVNNGN
11AA_SEQUENCE 1.0
P1:T01083 - hypothetical protein T10P11.5 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 14-May-1999
C:Accession: T01083
R:Kaplan, N.; Johnson, D.; Schutz, K.; Gnoj, L.; Hoffman, J.; Tili, S.; de la Bastide, M.; Granat, S.; Hamed, A.; Gottesman, T.; Hasegawa, A.; Shohdy, N.; Parnell, L.; Dedhia, N.; Johnson, A.F.; Lodhi, M.; Martienssen, R.; Chen, E.Y.; Wilson, R.; McCombie, W.R.
submitted to the EMBL Data Library, November 1998
A:Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.
A:Reference number: Z14248
A:Accession: T01083
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3449 <KAP>

A;Cross-references: EMBL:AC002330; NID:g2262135; PID:g3892055
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 62/2; 86/3; 115/2; 138/3; 157/3; 198/3; 262/3; 284/3; 323/1; 726/3;
960/1; 994/2; 1221/3; 1316/1; 1868/2; 2858/2; 2961/2; 3063/3; 3112/3; 3427/3
A;Note: T10P11.5

T01083 Length: 3449 May 30, 2002 09:26 Type: P Check: 8548 ..

1 MKWGTLIKDL KDKVGAETT ADLAGEAIS DPTPPSSSQ ASPSSSFAAL
51 AQHDENLISP TSDKIKLEL DKRYWEEFR SSSSEOEKEA ALNLVNTFC
101 RLVKQHANVD QLVTLMEPH IFSFVIGRAF VADVERLKDG SSHGANLLTA
151 IEVLASGPPD KQSLDSGIL CCLHTEFNAF LTVSASEGE KTVNYEKEVE
201 GSVVNIMKAL ASHPSAAQSL IEDDSLQLEF KMWANGSIMA FSRKVGGLVS
251 FHNIOHLKNA MOILGILLVN DNGSTASYIR KHLIKVLLM AVKDEPDPCG
301 DSAYTVGIYD LLECEVELSY RPEYGVRLK DDIRNAHYH FLYOFALLLS
351 SMPKDIYFAF DHSSPHKNRG SSDKKQPP LSKTRQNDSD EKQSLSLNS
401 RONDEFALKH FSPALSRLLD VLVTLAQTP IESSGTSTSL LSQTKLTGYS
451 RRQTPSANNR YDEPCEOGSG KVKDEAVQM LQDIFLKAEN KDLOAEVLNR
501 MKFIETSHLE NYRICOELKT VPLVLNMGG FPSSLOELIL KILEYAVTVV
551 NCVPEOELLS LCFLLQOPID SELKHTLSF FVKLTSPDQO YKKVLGEVGV
601 LEVLQDDLKQ HKLLRGPQY SGVSNHLDV PSSPSFKOHL DSQDAIISSP
651 KLMESGSGKL PIFEVERTIT VGWDCMISLL KNSQVNOEAF RSANGVTVIL
701 PFLIADEHRT SILRIFSCLI TGDIKOVHHE ELEALIDVLK SGMVTRVSGD
751 QYKLHYEVRD DIMGALMRIY GVNGSAQRFV GEATGRSLLL TTLHTFOGEE
801 ECRDESHLV YIKLFKHLR LITTAVCENA INRMKLSVI TSQTFYDILLV
851 ESGLLCYDLE RHVIOELLEL ALEVIVPPFL TSESMAEAM AECEKASFLV
901 KTASGOFNPD KOKIYNAGAV RVLIRSLILC TPKIQLEFLN LLERLARASP
951 FNKETLTSAG CVELLEIITY PFLQSSPFL SHALKIVEVL GAVRLSPSEL
1001 KMLCRYWQM RVNNGSPSLI GMMEKILME EDTGLECVSL APFVEMDSK
1051 TGHASVQVSL GERSWPPAAG YSFVCWVQFR NFLTTQELLES EYVKAAGSSK
1101 TPILSGOQSE ONIFRIFSVN AISNGSPSYA ELYEQEDGIL TLATSNSNSL
1151 SFSGLETEEG KWHHLAVVHS KPNALAGLFQ ASVAVYYIDG KLRHMGKLG Y
1201 SPSPVGKSLQ VIIGTSATCA RACGDSMAI LDLDTDMSS GIOKFEDSNR
1251 QGDSKAHC SG IWMIDRLGN LSIQLPGKKL IFAFDGTQCE FMRATGSFSL
1301 VNLVDELSAA ASLIGIPIRE GRIVGNVSLC RONVIGNSIR PYGMAVVLA
1351 LVEAESRDM LHMALSLLAG ALHONSQNVK DMETTYGYHL LALFLRPKMA
1401 LFDMOCLEIF FOISACEAF SEPKLESQ TTISMSPTET IPENNYEDPT
1451 LCKFOYETSS VSGHGDMDDF SGRKDSFSHL SELEMGDNPV ETSNCIVLSN
1501 ADMVEHVLDD WTLWTAIPVS IQIASIGFLE NLISILMYRS HNLAILROIN
1551 LVKHLVLTLQ RGDVEVLVLE KLVILLRCIL ENGFLTPELE DVVRAIMTF

1601 NPPEIKSQNS SMRESMGKHV IVRNLVLEML IDLOVITIKAE ELLEQWHKTV
1651 SSKLITYFLD GAVHPSSMRW IMTLGVCILT SSPNFSLEKEF ASGGQGLVR
1701 VLQSFYDSDP IYYILFCLIF GKPVPYPRPE VRMLDFRALM PDDGSIVELN
1751 FVDLLDSVVA MAKSTFDRLI MQSMLAHQSG NLSQVSARCV AELVEGYADM
1801 TGELOGKALM HKTYAARLMG GEASAPATAT SVIRFAYDLA KMCPOESAAC
1851 KNTFELQKCA DLYFSCVRAF HAVKIAQOLS MKAEQNITG GDDSSVEGNE
1901 CRVSHQDMST KTSISAGSFP QDQTSVTSYV DMYIPSDYVA VDKVENFLTT
1951 PPGESNKSFO GREYIAQODG DHVGSVASASS EMKSLDLTGS SSQVDPIDSR
2001 SSESFSMLES PLSEKSSLE VPFIPIPSKKS STISTPHPSH ISVSEFDASS
2051 DQSSGSQSS AVHTLFTTISP KVLLETDESG YGGGPCSAGA SAVLDPMAEV
2101 CADIMTEQIK AVQALESILE MLPLYVDEPC VVFQIGLCLS RVANYLERRE
2151 LRDEEDDKK LDKRKWSANL DAFQCMIVDR VYMGAFPOPT GVLRTLEFLL
2201 SILQANKDG RVEEVTSSGK GLLSIGRATR QLDAYVHSIL KNTNRTILYC
2251 FLPSFLITIG EEDLPRLGL LVESTKQOTS KLSGKESGID VSAVLQLLVA
2301 NKNIILCPNS LDTDLNCCIC VNLISLHDQ RKNVQNMASN IIKYLLVHRK
2351 SALEDLLYKK PHRGQKFDVL HGGFDRLITG NLPEFSKWLE SSEQIITKVL
2401 EOGAAVMWIQ YIAGSAKFPD VRMKGMDGRR TREMGRLRD TSKLDLKHWE
2451 QVNERRIALE VVRDAMSDEL RVVRQNKYGL ILHAESVPT HLQOLVHERG
2501 IFPMRISHGV EDLKWQLCPI EGPYRMRRKL ERCKLIKIDSL HNLLEGKLEL
2551 GEIELLSKKS EDGLVISMD SEPAFLISEL YSESFSEBAD DLKDVPSARN
2601 GWNNDRATST NAASLHNSLS FGGKSSSTAV SVPISVNTDE KSETGSPIKS
2651 SSGKMEIKH VEESEKELK DDGEYLIRPY LEHLEKIRER YNCERVGLD
2701 KHDGIFLIGE LCLYIENFY IDDHGICICEK ECEDELSIID QAQGIKKQFH
2751 GSLESKSKSS TLMSTTIKIG AVGGRAWAYG GGAWGKEKVR VTGNLPHPMW
2801 MKKLDVSHET LKRDYELRRV AVEIFSMDGC NDLVFPHKKE REEVFRNLIA
2851 MNLPRNSMLD TTISGSAKOE SKEGSRLEKL MAKSFTKRMO NGEISNFQYL
2901 MHLNTIAGRG YSDLTQYPVF PWILADYDGE SLDLSDPNNE RKLDPMGCO
2951 TPGESEEFRR RYESWDDPEV POFHYGSHYS SAGIVLFYLI RLPPESAENO
3001 KLOGGQPDHA DRLFNSIRET WLSAAGKNT SDVKELIPEF FYMPEFLENR
3051 ENLDLGERQOS GDKVGDVILP PWARGSVREE IRKHREALES DYVSENLHHM
3101 IDLIFGHQR GKAENAVNV FYHYTYEGNV DVDAVTDPM KASILAQINH
3151 FGQTPKQLEQ KPHVKRRPDR KVPHPRLKHS MHLVPRNIRK CSSSINOIIT
3201 FNDKLLLTGA NCLLKPRGYK KYIRWGFPDR TLREMSYDQD KILSTHENLH
3251 EGNQIQACAGV SHDRIYVVG AEDGLVSWR VSKDGPRGSR RLRLKSLCA
3301 HTAKVICLRV SOPYMIASS SDDCTVIIMD LSLSEFVRQL PNFSEVPTVV
3351 YINDLTGEIV TAAGSVLAWW SINGDCLSVV NTSQLEPTDLI VSVAGSTFSD

3401 WLEETWYVTG HOSGALKVWR MVHCTDPIKC CYRMQKQOSE TESRREDNN
11AA_SEQUENCE 1.0
PI:T48284 - hypothetical protein T22P11.220 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48284
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.;
Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A:Reference number: 224490
A:Accession: T48284
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-428 <BEV>
A:Cross-references: EMBL:AL162971
A:Experimental source: cultivar Columbia; BAC clone T22P11
C:Genetics:
A:Map position: 5
A:Note: T22P11.220

T48284 Length: 428 May 30, 2002 09:26 Type: P Check: 3770 ..
1 MTILPFLAAV FVLQLLSTLT VAEIKSFTIS NDSRPVILL KFGIIEIGHV
51 TVSVSSSVSVL SPILDSKLG FEVLSEESLP HVLLLEQNF SFCVLDSHYI
101 LHFFTFVDLS PPPRSQFSKS YPITSPNDYS LPFANCVPET RVSMKVHTEI
151 YHDLYPNGSR DYLLAGSAQL PGLYLVEFLC YLSFLCFWLC FGMNKKQIVK
201 RIHLMTALL LVKSLTLICA AVYKHVVKVT GRAHGNIVF YIFQFISVVL
251 LFNVIVLIGN GWSFLKPKLH VKEKKLLIV VPLQVLANIA SIIVGETGPY
301 TQDWVSWNOI FFLADITCCC AIYFAMVWSM CCLRETSKTD GKAVKNLAKL
351 PVLRRFYVLV IGYLFFTRIV VVVMKKKADF TYQWVSNAE EIALTSFYCL
401 MFYMERPIEK NEYCDVDDEE EIVELSLK

11AA_SEQUENCE 1.0
PI:T39938 - hypothetical protein SPBC23E6.04c - fission yeast
(Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T39938
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Pohl, T.
submitted to the EMBL Data Library, May 1998
A:Reference number: 221892
A:Accession: T39938
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1649 <MOO>
A:Cross-references: EMBL:AL023287; PIDN:CAA18872.1; GSPDB:GN00067;
SPDB:SPBC23E6.04c
A:Experimental source: strain 972h-; cosmid c23E6
C:Genetics:
A:Gene: SPDB:SPBC23E6.04c
A:Map position: 2

T39938 Length: 1649 May 30, 2002 09:26 Type: P Check: 5030 ..
1 MASSIQKQLK NIOSNNVLKI NKIRRAPSL YDPKVAADMD LEEIYTVAS
51 GFHELAVHEP RLTYFEKTL GEGSQVQDRV LLNRTENEKI DLCEVQILRL
101 LAPFTEKNA LKVLLEMLIR FSIHEYVSDE FILSFLPFHD HPFARILGC
151 SKPKSRPLLF LENAIKMPVT LSRADIVHAL SRDKEFFAMF AQFVQNTAES
201 HNMYPELARF WAGTMEVLV AMHSSNEDPN VLLDRFLRV SYAVSYVSSI

251 DFQIAGFMLL SSIAASUPLS PSIIPLVSA ITDRUSEONI KPALICVGH
301 LQFCSSFEED HEQLEKLESF GASSLLIELS QEHRLEDEFV SYWVSLIKSR
351 KQKDKRRLIS LLDFTSISQIR VTHEQAKELL SVIPVNDQFK ALQSYRIL
401 SVIQPERKEG KLDNLINTLQ DKRKSSTFSK KDREYLLKRI SEIDSQTSFE
451 QCLAYADSAAL DLDSSVFISL LSKFGDKIPF LFLCIANGSE RIILLSLIEL
501 RKTIEENKDV DYQIILPVVL YSIQSKDTEV RSRALNLILT FLELRNENLE
551 FSIIYGMDDN DNKNLRMLSP VETKYCCSDL LLDSSSEIGL DGTLYFSYIP
601 ERLFTEKKPK NASKEIAVTS FLSSHAACSK LSNRVLLLE ILFRVHGKVE
651 DAKMQLLPR LEQLSEENSE KFKIVSKREV EALVNCFNHT SFTSLISFLS
701 SNIVLSQAIC RRIVEIQSHL KDPQRLFEVK AVISQDEQPH YYVDYLD
751 IPDTYFRRKLI GSVRLVKEKN PALAKRRRID SHIPDCDVOR LTRILELLET
801 KNAASYPKLA SPLFEVLNSV IALKEDIVSS NYLLOLLGL LYEMIGASPI
851 TELSPSIRID TLVGCIRSTN NPQIQNKALL LVSALANAAP EAVLHGVMPI
901 FTFMGSTIVLS RDDAFSIVHI EQYKTVISA LIRLQKDFDS SLVSCFVNA
951 FPHIPQHRRL RLRYLYQTI GSNRFLSVVL IQFAEKMLLA KSTNVVAIHD
1001 FCLTIVSFS VADRIQISINQ CSRFLKSL EQSNGSDNGK AVSLIKDEL
1051 PMDVLATLG SLRVKYLELI SLVSKAKNFA FDLAKIMENS VDSFEIQAG
1101 LFESIKLLIT LSQSSNEME LGHYVALRS VIHLLPNELF CTVLGKLLHD
1151 ERALLRRKAL SIVQQRVQOG SKVSALTALI PDVTYNISNY SDEETTQOLAM
1201 DCLAVMAKRF SASPELFISP IEVVGSPYGL KNSARDVQS AIVCTIVLTN
1251 TLAARILPYL ADIVNYSLSI LDDARKDPEG DLLELACFSM MIDFEKVLPE
1301 FSSSYVEPTI KCALASDBAF EHDAGIELLE ETIANPIPTR LLMKSIFAAM
1351 PECARLGSTA ALRLLELIEL ALQNSSRSAL GTVYKSLFKF FLDSFDSRRS
1401 LLEAEDVDNV ETQAVNVFLK FVMKLSDTTF RPLFLHLSW ALEDLEYETDP
1451 SGIVSRQTFE YNFLTIFLDT LKSIYVNYA YVLDITIELL SSKPTNSEVR
1501 HLVNSSLVSA FENDTEFTWM VPARFGKISP VLIEQIQYAP LLDKVLVKA
1551 IVELASVASS SDNFRSMNTQ LLQYLRSNI NARLLAIQIQ TQLYGRGEN
1601 WISTLPQVVP FIAELMEDDD DQVETATDEL VRIIDRLGE NESIQDIYLT

11AA_SEQUENCE 1.0
PI:S36335 - U2 snRNP 40K protein - Trypanosoma brucei
C:Species: Trypanosoma brucei
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C:Accession: S36335; S28637
R:Cross, M.; Wieland, B.; Palfi, Z.; Guenzl, A.; Roethlisberger, U.; Lahm,
H.W.; Binderelf, A.
EMBO J. 12, 1239-1248, 1993
A:Title: The trans-spliceosomal U2 snRNP protein 40K of Trypanosoma brucei:
cloning and analysis of functional domains reveals homology to a mammalian
snRNP protein.
A:Reference number: S36335; MUID:93209224
A:Accession: S36335
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-326 <CRO>
A:Cross-references: EMBL:X69934; NID:g10563; PID:g10564

A:Note: part of this sequence was confirmed by protein sequencing

S36335 Length: 326 May 30, 2002 09:26 Type: P Check: 3173 ..

1 MRLTLDITRR APQFTNALRQ RELDLRGIGI TVLEHTLTF LNDSPDYNL
51 SQNPLARLEY FPGDSAPLAT AAAQNSSAKP ASRMMLRIQT LVVHRNLTH
101 VSEATCAAVL PNLRAFAVDH NEFRELRLDL FLSHWKKLEI LSEHNPIITI
151 SEDNARLRAY VVELCPITKL VNYQRTQVD RQNVETMRKE FVGLVEGWR
201 LEAKQLLOON SAPTEADASA SESVKKIKRR SRHAREASK NGSADTAPEP
251 AASKVETPAT PSVAAEAGEE EASNALQAR LEALEEKMAA AETEEELMEL
301 QQELTELETI MKHQASKGT KKTRTS

11AA_SEQUENCE 1.0

PI:T19297 - hypothetical protein C15A7.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T19297

R:McMurray, A.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19104

A:Accession: T19297

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-814 <WIL>

A:Cross-references: EMBL:Z67735; PIDN:CAA91530.1; GSPDB:GN00028; CESP:C15A7.1

A:Experimental source: clone C15A7

C:Genetics:

A:Gene: CESP:C15A7.1

A:Map position: X

A:Introns: 34/3; 84/3; 191/2; 243/2; 279/2; 325/1; 396/1; 423/2; 476/1; 548/2; 611/2; 663/2; 698/2; 744/1

T19297 Length: 814 May 30, 2002 09:26 Type: P Check: 7776 ..

1 MIEPLMIVL ENTIPIRYSE ADESESRHQ YLDHRKLED VLGLDGSMP
51 SELPVEIGAQ LKLDHISRD EKTGFLTAL TKFMSDER LOWDPHSYSG
101 FTRRLFEVD FWLYQIWKR IYMTNTYERR QSKTIDLSN TLVEMDIHSK
151 GFVMTTTRIL LKTECFLEDEK GYPHDFQNS FSEIPNMNAN DIFGSSMLK
201 AKFTKDFDTP QQVVRVQDEQ INDVTEEDLY LEMNRIYVLD LITREPSLVR
251 SFVSEFIVLQ RKNVYVHLQ YAPMIFICTC LIISGFIPSE YALPLIVENL
301 ISEYLYFHHI KDVLPDDEG TPTIALIATF TLAETMSIIG WKMFIITIMT
351 KKQKAYFLNS SKFPLDSGEF NIRVLRRLLT IDKFLMYMLI AQSGKPTTI
401 FLQLFFQRA ELRMLTLKVN RQNFPPVLAS QYSVKINISF SLEESSYNS
451 YFHCAARYPS SCGVTMGNGK KLHGFDHGHG FMSVINAEL VSYLYQPE
501 SFIFQWIDE RLKWNPNAYS GYREIVEKTF EFRKDNNCWM PIVKYSYDR
551 RYSELDLLEF SDARTLISYK GEIKTALQTM VTTCQFSEG EYPNDYQNS
601 IMLIPQONAD EFRFVSPNGY CNPKFENLEH RAVRVHDLHL MGVESNIEYT
651 FANTYFPAEE VTGFVYMAKT EFRFNLMEKR VNKLNVKLS IPSILISMEL
701 ILAGLPNGY SIFGYACEF VEVMHGLYVS KILPNDIRGI PYHGALALCF
751 LIETVFLFCW KVESVARNQ KLFAQLPNVF ADNRVEVLKE VFYVDRVVCV
801 ALCLQFLSVI YKTR

11AA_SEQUENCE 1.0

PI:T23772 - hypothetical protein M162.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T23772

R:Basham, V.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19797

A:Accession: T23772

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-170 <WIL>

A:Cross-references: EMBL:Z82278; PIDN:CAB05254.1; GSPDB:GN00023; CESP:M162.2

A:Experimental source: clone M162

C:Genetics:

A:Gene: CESP:M162.2

A:Map position: 5

A:Introns: 75/1

T23772 Length: 170 May 30, 2002 09:26 Type: P Check: 6239 ..

1 MFPKLFLAI LIPSLTDCP DIDGTREIK GICYKFVTSI MKYEDARDWC
51 HHHNPAGPSW LAYVPDQETS NFLAVYAGSI FEGEGKHFWI GLNRDPISKT
101 LSWDTGLSVS YTNFGSNVAQ NYFSENTINT KWNTIGDDEV HNFVCSYRPS
151 TVPATVTRQP QARRLAAMKN

11AA_SEQUENCE 1.0

PI:T26443 - hypothetical protein Y113G7B.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26443

R:Lennard, N.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20215

A:Accession: T26443

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-331 <WIL>

A:Cross-references: EMBL:AL110477; NID:el542121; PIDN:CAB54331.1; CESP:Y113G7B.9

A:Experimental source: clone Y113G7B

C:Genetics:

A:Gene: CESP:Y113G7B.9

A:Introns: 32/2; 58/1; 178/3; 212/3; 247/3; 285/2

T26443 Length: 331 May 30, 2002 09:26 Type: P Check: 4393 ..

1 MEITIQLLCL LGIISAITTI LLNINLVKI VLNPSKRKND MYLFYFRFTL
51 DIFFGAGLFS YIAYTLNME APEMFQYRS LIVLLALPWS HISTCRSIIA
101 LSISIDRSIA TCFPIYFEKN RKKIPNMPVL LIGSLGLAE EYMLFGCSY
151 NMEIPKTCIV FGCAFNQCFE HWLIQRSII FSLIVLFSLI LSIKLMMNS
201 VKHQSNNOI SKANRLALLD TCTVLLFDFL PAFCGHMPPT APMEFERNVG
251 SYNPLKITG CAIESTVTR VLLFRTPENS ASTPVOKLTY TLRFLSLAQ
301 FPFIFKNIFE HFLRSRAVTV LYKNILKLIQ F

11AA_SEQUENCE 1.0

PI:H83904 - hypothetical protein BH2040 [imported] - Bacillus halodurans

(strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: H83904

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii,

F.; Hirama, C.; Nakamura, Y.; Ogasawara, N.; Kuhara, S.; Horikoshi, K.

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*.
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: H83904
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-818 <STO>
A:Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BAB05759.1; GSPDB:GN00137
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2040

H83904 Length: 818 May 30, 2002 09:26 Type: P Check: 9494 ..

```
1  MTTKFINELK NDLPLDQGEWC FEPHMLRSA GEPKMLTQF TETDTMKLVE
51  EKKRLDEKWN ETISECNNEL EKMLDLTTKQ KKLISOLRR YKIPTAKGST
101  ELLANLGAGF QIKLEDMEIK MNYLEAQISK TFQTELDQRQ HHLKQVYLEN
151  ESLQEAFLS SPDAYHNGLK HFLSKELEQS RPRGORQEQ LAAIYLQRFY
201  SKNETASYYG PTNYGTFTSQ PGLLELHVNG PIKROLFSY WAVQSLANVM
251  AEDDSILPYL KPPLSPFIKR ENSNLROSSN GKLIQLPEIY HKMIRYSDGN
301  HNIHQIAAML DLSTAECIAR IQKLEKRII LLEIDPFTSV FDPFEKLEK
351  IQQLPNDCSS KSKWLHVINE WSSHLLTWKE SESFQQRHI LTLLENSFEQ
401  NVGLSSRKDR GKHYTDRLIV FEEARGDVET CRIGLNMOKQ WKDQLEPIFK
451  LISCRAVEH RALTEVAYKE YQSFNSDVNF LTFALHMQGC KTEALQYAKE
501  KLEFEDNRLS EWTALHPFEK GSIVIAKEDV DAFCLNFPQD DVALFSPDIM
551  IAADDVDAIN EGKYKLIGE VHSGIQWSV LNSVYPDQLR LNEEIIYHHLG
601  PTLQSFMLEH VQPRAPGKTF RPELSQGTIV ENLGRSMKSR EYVRSIAELG
651  LIYENERFYV TEGEKPKLMD LETDVEPLNQ IFSLPSVKSF SIQIGENTPR
701  IEINGVVFQR ERWTFACKEL LVRVNGYTDW LLDWATDLR KNYNMPRYVY
751  ARGNNPEKPI FVDHFNEFTL EVLVQLLKRN EHVSITEWVP AQDSLWFTRG
801  ENKHTAEFRF SVIHENLT
```

!!AA_SEQUENCE 1.0
P1:B84135 - RNA polymerase ECF-type sigma factor BH3882 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: B84135
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirama, C.; Nakamura, Y.; Ogasawara, N.; Kuhara, S.; Horikoshi, K.
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*.
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: B84135
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-154 <STO>
A:Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07601.1; GSPDB:GN00137
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3882

B84135 Length: 154 May 30, 2002 09:26 Type: P Check: 5308 ..

```
1  MKVSELYEEL KQDLYRFART ISRNEQAND LIQDAMEKSL RCTELLDWPR
51  HKQKWFYRV MKNQLIDIRR KEREEDWDD SVYPPFTTAG TSYVEMVDLL
101  SKLPASLSDL VFKEFWLGLT SQEISEKGVV PASTVRYRIR LAVKKLREYM
151  KEEG
```

!!AA_SEQUENCE 1.0
P1:D97337 - mismatch repair protein Muts-Like Appase [imported] - *Clostridium acetobutylicum*
C:Species: *Clostridium acetobutylicum*
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: D97337
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, H.M.; Dubois, J.; Qiu, D.; Hittl, J.; Wolf, Y.I.; Tatusov, R.L.; Sabathe, F.; Doucette-Stamm, L.; Soucaille, P.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum*.
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D97337
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-577 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK81487.1; PID:g15026659; GSPDB:GN00168
A:Experimental source: *Clostridium acetobutylicum* ATCC824
C:Genetics:
A:Gene: CAC3563

D97337 Length: 577 May 30, 2002 09:26 Type: P Check: 4438 ..

```
1  MDEDKLYMYS AIATLALGIS AFVFPQLAIG YKNYIYELGT ALSTIAIFLA
51  FYMKAKIHK LNYERKKRYW GEFQKKRNT RNVKKEFEFH KEDNADEFNV
101  DEQTFEDLNM NKVFEADRT LTPSGEMLY KIFKPEFSK EKLERNSTII
151  KIFQNKREVR EKVGLLEIKL GRKKENGVTD LMKDIEVDY KYKYVENFLF
201  WGTLASILTI PILKYGIIL LAAFILMNTI EHNKFNKVE LYYQSLGYLN
251  GIINTANRIS KINCPKIKYV TDSLKATSSK LMKVAKKTAG IERYEGVDII
301  GDALYNILPI EERKFNNSIN DIRKLKELK ELYKALGEID ALMSIASFRQ
351  WIQYCEPEEF VDNGRITKS EIIHPLENP VNSNIDLDGK GIMLTGTNMA
401  GKSTFLRTIG LNSLQAOTIY TCAAKYKTS FFKIMTSISP EDNITSSGKSY
451  YFREAEALKR IINQCSDDRP VLCIIDEIFR GTNPIERVNA SAEILNYIEK
501  HNTLTIVATH DLELTEILKE DYLCYFSED IDDEGLQFDY KIKYGICKTR
551  NAVKLLKYLE YPNEITQKTN DRLAKIL
```

!!AA_SEQUENCE 1.0
P1:A10020 - probable membrane receptor protein YPO0164 [imported] - *Yersinia pestis* (strain CO92)
C:Species: *Yersinia pestis*
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: A10020
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; Sebahia, M.; James, K.D.; Churcher, C.; Mungall, K.L.; Baker, S.; Basham, D.; Bentley, S.D.; Brooks, K.; Cerdano-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Feltham, T.; Hamlin, N.; Holroyd, S.; Jagels, K.; Leather, S.; Karlyshev, A.V.; Moule, S.; Oyston, P.C.F.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.G.
Nature 413, 523-527, 2001
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AI0020
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-686 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC89027.1; PID:q15978269; GSPDB:GN00175
C:Genetics:
A:Gene: YP00164

AI0020 Length: 686 May 30, 2002 09:26 Type: P Check: 7936 ..

```
1 MISGKNKVCK KLLACAVTS AISAFASNAI ASNAIADTTS LAAMNRSAIN
51 ISSTHNSNKK LNDADVITVT APLYSPLVVV TSPKVPROPV PASDGTDLR
101 TIPGFSLVRN GGTNGDVFR GMFGSRLKIL TDGAEILGTC PSRMDAPTST
151 ISPEFDLLT VYKGPQSVLW GPGVSAGTIL FERERPHFDQ SGVKGNASVL
201 AGSNGRWCKN LDASLGNQCG YLRMGNSQR AADYKDGNNV RVPQWDKWN
251 ADMALGWTPD TDTLLELSVG KNGGEARYCG RMDGSQFLR ESLARVERKS
301 NIGEVLDKIE AKINTTYVDH IMDNKTLRTP PMAMSSNLD RMLGGRVMG
351 TWLMQDLKLD AGTDMQTNTH RKNQGNWDK NARFMNNGVF GELTWAASEQ
401 NKLISGARLD RHQVTNYTRA GEPTRSATLP AAFMRVEHNL ANAPVMTYAG
451 LGHTERFPDY WELFSPKFGP AGSTSAFEGV KSEKTTQIDI GGQYSGEQFR
501 GWLSAYIGRV NDFILFKYDP KNLRIQADN INATIMGEM GMEYALTPYW
551 KADASLAYAW GENTRDNRLP PQIPPLEARL GLTFEAGDWS SSALMRLVNS
601 QHRVAINEGN VVGKDESSSA GFGVLSANVA YKVNQVKLS SGIDNLNKA
651 YSEHLNLACN SGFGYSVNTP LNEPRTLWA KLNLT
```

1|AA_SEQUENCE 1.0

PI:AB1288 - transport proteins homolog lmo1706 [imported] - *Listeria* monocytogenes (strain EGD-e)

C:Species: *Listeria monocytogenes*

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AB1288

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1288

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-289 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAC99784.1; PID:g16411160; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo1706

AB1288 Length: 289 May 30, 2002 09:26 Type: P Check: 1598 ..

```
1 MMKKVVKYK QNGIVQVGA VSARVGRNDV SGNAQLAYY MLFSIFPML
51 IAAITLALYLH IDKDSVFNMI KEFAPDQIMD FLEENLNL TLKNGGLLSI
101 GIATLWSAS NGMNAVMSL NKAYGVTKR NYVQRLISM FFLAMLATV
```

```
151 GATLLLVFG QQIGMFLINH LNSEDFLSF WNNLRWTVTL IVIEVVFTEL
201 YWVAPNRST LISVLGALF STIGWTVASV GFAYVNNFG NYSATYGSIG
251 VITILMLFY LTGIILMIGG ELNATLAIRK KKELEGIN
```

1|AA_SEQUENCE 1.0

PI:A1659 - transport protein homolog lin1818 [imported] - *Listeria innocua* (strain Clp11262)

C:Species: *Listeria innocua*

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: A11659

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: A11659

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-289 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97049.1; PID:g16414320; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lin1818

A11659 Length: 289 May 30, 2002 09:26 Type: P Check: 1480 ..

```
1 MMKKVVKYK HNGIVQVGT VSARVGRNDV SGNAQLAYY MLFSIFPML
51 IAAITLALYLH IDKDSVFNMI KEFAPDQIMD FLEENLNL TLKNGGLLSI
101 GIATLWSAS NGMNAVMSL NKAYGVTKR NYVQRLISM FFLAMLATV
151 GATLLLVFG QQIGMFLINH LNSEDFLSF WNNLRWTVTL VVIEVVFTEL
201 YWVAPNRST LISVLGALF STIGWTVASV GFAYVNNFG NYSATYGSIG
251 VITILMLFY LTGIILMIGG ELNATLAIRK KKELEGIN
```

1|AA_SEQUENCE 1.0

PI:AC2420 - hypothetical protein alr4915 [imported] - *Anabaena* sp. (strain PCC 7120)

C:Species: *Anabaena* sp.

A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AC2420

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120.

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AC2420

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-313 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAW7614.1; PID:g17134053; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr4915

AC2420 length: 313 May 30, 2002 09:26 Type: P Check: 9610 ..
 1 MKNLDYCCQS FSQLNVSSSTR KRGNAQYKPI LVLTVIDLIT QGLITDNGIR
 51 VSEELIQTFY RYWSILGSES YKGLHYPPF HLQSDGFVHL VEKPEFNGIQ
 101 PKTTNKLKEA VEYAYLDKEL FNLQDESPR QELIDALVTT FFOEQODELE
 151 EMLQINQSFQ DADLATETSS GSINLDNPK WGCRAKAIIRN AFEKTIIVRY
 201 YDYKCAFCGL RYTKAINQNT VDGAIKPPA QFYDSRIHNG IALCKNHMA
 251 FDRGWFTVDE QYKIIVSKEL QEISPHSKPM KDFHGERLLI PNOEQYLPEL
 301 ESLQWHRQNV FOA

!!AA_SEQUENCE 1.0
 P1:AF2463 - CAB/ELIP/HLIP superfamily of protein [Imported] - Anabaena sp.
 (strain PCC 7120)
 C:Species: Anabaena sp.
 A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AF2463
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120.
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AF2463
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-59 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA076961.1; PID:g17134401; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: asr5262

AF2463 length: 59 May 30, 2002 09:26 Type: P Check: 4158 ..
 1 MDTTKISAS VVEDRNSWRN GFTPQAEIYN GRLAMIGFLA ATLIELFSGQ
 51 GFLHFWGIL

!!AA_SEQUENCE 1.0
 P1:AG2271 - CAB/ELIP/HLIP superfamily [Imported] - Anabaena sp. (strain PCC 7120)
 C:Species: Anabaena sp.
 A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AG2271
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120.
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AG2271
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-59 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA075425.1; PID:g17132860; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: asl13726

AG2271 length: 59 May 30, 2002 09:26 Type: P Check: 4679 ..
 1 MSGFKNPKST VTSPNAVRF GFTSESESWN GRFAMIGFLS IYVLEAFSGQ
 51 GFLHFWGIL

!!AA_SEQUENCE 1.0
 P1:AH1862 - CAB/ELIP/HLIP-related protein [Imported] - Anabaena sp. (strain PCC 7120)
 C:Species: Anabaena sp.
 A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AH1862
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120.
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AH1862
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-40 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA072407.1; PID:g17129794; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: asl0449

AH1862 length: 40 May 30, 2002 09:26 Type: P Check: 580 ..
 1 MGFNHQSESW NGRAMIGFL AATAIEFFSG QGFLHFWNIL

This Page Blank (uspio)

```
1 FINDPATTERNS on swp: * allowing 0 mismatches
1 1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
2A5E_MOUSE ck: 8649 len: 387 1 Q61151 m serine/threonine protein phosphatase
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
106: LQLVY EFFIRFWESQ EEPQPS
BRAC_HALRO ck: 462 len: 471 1 P56158 halocynthia roretzi (sea squirt). brain
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
103: VHPDS PNEGSHWMKQ PVSFN
CHS1_TUBUN ck: 8541 len: 189 1 P55003 tuber uncinctum (burgundy truffle).
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
123: LNSHR CIFQRFWSTL NPNVC
DNAJ_SYNP7 ck: 2256 len: 287 1 P50026 synechococcus sp. (strain pcc 7942)
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
68: RRAEY QRFSTRWQQQ GAASV
IRF3_CHICK ck: 3644 len: 491 1 Q90643 gallus gallus (chicken). interferon
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
446: LVKLV POFCEYWEQ VQRCG
PKN1_CHLMU ck: 1219 len: 934 1 Q9PK92 chlamydia muridarum. probable serine
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
664: KSSKM ELFLSFWTGF TIFLP
PKN1_CHLPN ck: 8398 len: 932 1 Q92986 chlamydia pneumoniae (chlamydophila)
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
662: RSSKM ELFLSYWSGF IPHLN
PKN1_CHLTR ck: 4900 len: 934 1 O84303 chlamydia trachomatis. probable serine
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
664: RSSKM ELFLSFWTGF SLFLP
PUCC_RHOSH ck: 7732 len: 459 1 Q02443 rhodobacter sphaeroides (rhodospseudomonas)
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
244: KPEKE PEFGDHWRER ISREN
TCE2_AVEA ck: 4989 len: 535 1 P54411 avena sativa (oat). t-complex protein
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
3: MA LDFDEYWRPF IILRE
```

```
1 TRL2_HUMAN ck: 5055 len: 1,503 1 O94759 homo sapiens (human). long transit
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
862: MKKAA LYFSDFWNKL DYGAII
VE5_RHPV1 ck: 8743 len: 157 1 P24834 rhesus papillomavirus type 1 (rhph)
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
37: YACKP PPFCCEFWLCF CCCFC
VP41_BPAPS ck: 4265 len: 460 1 Q911q7 bacteriophage apse-1. putative protein
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
391: DCGNI LVFFSHWMDL EGYQQ
Y306_MYCPN ck: 3483 len: 395 1 P75343 mycoplasma pneumoniae. hypothetical
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
169: LEQNH EGFVSYWFSE AGMIF
YA48_MYCPN ck: 4533 len: 518 1 P75066 mycoplasma pneumoniae. hypothetical
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
251: SLWNF QEFISYWTGQ GKDVK
YC03_KLEPN ck: 2748 len: 504 1 Q48449 klebsiella pneumoniae. hypothetical
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
293: GEGRP QSFSSFWDGF TGHDN
YC34_MYCPN ck: 7095 len: 417 1 P75537 mycoplasma pneumoniae. hypothetical
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
164: SLTSA QKFKEYWLKQ GKDYT
YCF2_MESVI ck: 5630 len: 890 1 Q9MUP8 mesostigma viride. hypothetical 1
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
795: QAASI LDFNEFWTEQ DYPWQ
YCLF_BACSU ck: 9486 len: 492 1 P94408 bacillus subtilis. hypothetical t
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
28: RGLFT LFFTEFWERF SYYGM
YF11_MYCPN ck: 3241 len: 260 1 P75275 mycoplasma pneumoniae. hypothetical
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
2: M QAFKEYWQK KKDVT
YG44_SCHPO ck: 5030 len: 1,649 1 O60179 schizosaccharomyces pombe (fission)
```

1	337: EHRLD	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFVSYWVSL	IKSRK
1	YL92_YEAST	ck: 3922 len: 518 i P18634 saccharomyces cerevisiae (baker's yeast) (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(S)(H)WXX(F) PTFNHMQEF	NKIDN
1	146: NARLP	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(F)WXX(L) ESFYRFWAGL	AYDRI
1	YXIK_BACSU	ck: 7683 len: 153 i P42302 bacillus subtilis. hypothetical protein (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(F)WXX(L) ESFYRFWAGL	AYDRI
1	Q9X4B6	ck: 2369 len: 504 i Q9X4B6 escherichia coli. k30 capsule biosynthesis (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(S)(F)WXX(F) QSFSSEFWDGF	TGHDN
1	293: GEGRP	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(S)(F)WXX(L) QSFSSEFWDGF	TGHDN
1	Q9L1E1	ck: 7894 len: 70 i Q9L1E1 streptomyces coelicolor. sc3d11.21 protein (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(E)(F)WXX(L) LSFDEFWAHL	NK
1	59: DGDKQ	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(E)(F)WXX(L) LSFDEFWAHL	NK
1	P95529	ck: 7071 len: 492 i P95529 pseudomonas putida. tdnq protein (ec 1.11.1.1) (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(Y)WXX(L) EEFIRYWMHL	RRSEW
1	449: REAFG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(Y)WXX(L) EEFIRYWMHL	RRSEW
1	Q44310	ck: 4158 len: 59 i Q44310 anabaena sp. (strain pcc 7120). rna (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(H)(F)WXX(L) QGFLHFWGIL	
1	50: ELFSG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(H)(F)WXX(L) QGFLHFWGIL	
1	Q52209	ck: 5380 len: 324 i Q52209 pseudomonas putida. pupr protein. 12 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(R)(Y)WXX(L) QHETRYWVTL	KPRMA
1	310: LQLEV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(R)(Y)WXX(L) QHETRYWVTL	KPRMA
1	Q9F0G8	ck: 6799 len: 364 i Q9F0G8 rhizobium meliloti (sinorhizobium meliloti) (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(F)WXX(L) EAFADFWSKL	SLAFG
1	144: EAVPY	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(F)WXX(L) EAFADFWSKL	SLAFG
1	Q06954	ck: 8871 len: 460 i Q06954 vibrio cholerae. rfbg protein. 6/2000 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(F)WXX(F) EKFSFWDPF	VFEYN
1	158: HLLGG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(F)WXX(F) EKFSFWDPF	VFEYN
1	Q00620	ck: 5873 len: 683 i Q00620 pseudomonas stutzeri (pseudomonas putida) (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(D)(Y)WXX(F) Q00620	
1	402: GLGHI	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(D)(Y)WXX(F) Q00620	
1	Q93OE9	ck: 4054 len: 567 i Q93OE9 azotobacter vinelandii. phb synthase (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(S)(F)WXX(Q) PSFTSFWSAQ	TPFVA
1	5: MDQA	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(S)(F)WXX(Q) PSFTSFWSAQ	TPFVA
1	Q93N64	ck: 559 len: 314 i Q93N64 coxiella burnetii. hypothetical protein (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(Y)WXX(L) LDFVDYWFGL	RDSNY
1	164: LCCHL	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(Y)WXX(L) LDFVDYWFGL	RDSNY
1	Q9P981	ck: 4445 len: 385 i Q9P981 robillarda sp. (strain y-20). end (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(E)(F)WXX(L) EQFGFHWGEL	ASHFN
1	150: TAATT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(E)(F)WXX(L) EQFGFHWGEL	ASHFN
1	Q9UW21	ck: 7245 len: 429 i Q9UW21 candida albicans (yeast). hypothetical protein (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(C)(F)WXX(L) EVEFCYWDEL	PNNKS
1	128: NPVLG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(C)(F)WXX(L) EVEFCYWDEL	PNNKS
1	Q9UI23	ck: 6679 len: 190 i Q9UI23 homo sapiens (human). pro0529. 3/ (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (C)XFX(C)(F)WXX(L) CHFVCFWVCL	PLRVS
1	18: VYVCV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (C)XFX(C)(F)WXX(L) CHFVCFWVCL	PLRVS
1	Q9NZH3	ck: 2970 len: 280 i Q9NZH3 homo sapiens (human). c-type lectin (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFYSYWTGL	LRPDS
1	191: SOSYS	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFYSYWTGL	LRPDS
1	Q9H245	ck: 5880 len: 531 i Q9H245 homo sapiens (human). clorf28. 12 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PVFLRFWETL	DRYMW
1	509: RHLDR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PVFLRFWETL	DRYMW
1	Q9BXJ8	ck: 1303 len: 343 i Q9BXJ8 homo sapiens (human). transmembrane protein (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LFFGHFWOLF	NALTL
1	279: FLLPF	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LFFGHFWOLF	NALTL
1	Q9NV66	ck: 4753 len: 732 i Q9NV66 homo sapiens (human). cdna flj109 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQRF	LDLTK
1	547: KKIDR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQRF	LDLTK
1	Q9H9C4	ck: 6460 len: 561 i Q9H9C4 homo sapiens (human). cdna flj128 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQRF	LDLTK
1	376: KRIDR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQRF	LDLTK

1 Q9H5L7 ck: 1905 len: 355 1 Q9h5l7 homo sapiens (human). cdna: flj23316
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(P)XFX(R)(F)WXX(L)
333: RHLDR PVFLRFWETL DRYMV

1 Q96KN6 ck: 6589 len: 1,449 1 Q96kn6 homo sapiens (human). putative trp c
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(L)XFX(D)(F)WXX(L)
LYFSDEFWNKL DVGAI

1 Q96JM4 ck: 5997 len: 1,227 1 Q96jm4 homo sapiens (human). kiaa1801 prote
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(E)XFX(S)(Y)WXX(L)
EAFSSYWLPL LQNTT

1 Q96A07 ck: 1525 len: 355 1 Q96a07 homo sapiens (human). similar to hyp
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(P)XFX(R)(F)WXX(L)
PVFLRFWETL DRYMV

1 Q9V696 ck: 9003 len: 2,040 1 Q9v696 drosophila melanogaster (fruit fly).
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(E)XFX(D)(H)WXX(Q)
EHFSDHWMKQ NQDPF

1 Q9V9K3 ck: 7480 len: 337 1 Q9v9k3 drosophila melanogaster (fruit fly).
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(E)XFX(D)(H)WXX(Q)
EDFDEFWNVL RIVGF

1 Q9GN87 ck: 9284 len: 209 1 Q9gn87 anopheles dirus, and anopheles dirus
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(Q)XFX(D)(Y)WXX(Q)
QRFGDYWYPO IFAKO

1 Q9NL42 ck: 320 len: 491 1 Q9nl42 clona savigny1. brachyury. 12/2001
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(P)XFX(S)(H)WXX(Q)
PNFGSHWMKQ PVGFS

1 Q9BLF5 ck: 1820 len: 308 1 Q9blf5 olkopleura longicauda. brachyury pro
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(P)XFX(S)(H)WXX(Q)
PNFGSHWMKQ SIGFS

1 Q9BLF4 ck: 7868 len: 403 1 Q9blf4 olkopleura longicauda. brachyury pro
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(P)XFX(S)(H)WXX(Q)
PNFGSHWMKQ SIGFS

1 Q17041 ck: 8670 len: 491 1 Q17041 caenorhabditis elegans. hypothetical
(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(P)XFX(S)(H)WXX(Q)
PNFGSHWMKQ SIGFS

	275:	YSSVF	(L)xFx(S)(F)Wxx(L) LSFVSEFWIDL	KALPA
1	Q9XYXS1	ck: 646 len: 438	i Q9xyxs1 ciona intestinalis. brachyury pro (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(S)(H)Wxx(Q) PNFGSHWMKQ	PIGFS
1	IHPDS			
1	Q9GUZ8	ck: 9657 len: 334	i Q9guz8 oikopleura dioica. brachyury prot (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)xFx(S)(H)Wxx(Q) PNFGSHWMKQ	PLSFS
1	MHPDS			
1	Q9NHK6	ck: 7692 len: 110	i Q9nhk6 plasmodium falciparum. pfemp1 pro (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(E)(Y)Wxx(L) EKFREYWVAL	NRKEV
1	DKLSL			
1	Q9NHG3	ck: 7219 len: 110	i Q9nhg3 plasmodium falciparum. pfemp1 pro (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(E)(Y)Wxx(L) EQFREYWVAL	NRVOV
1	KKLLT			
1	Q9U2I8	ck: 2401 len: 569	i Q9u2i8 caenorhabditis elegans. y41c4a.2 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(E)(Y)Wxx(L) LKFIETWKSL	GVDVL
1	IESYK			
1	Q9U2X4	ck: 4393 len: 331	i Q9u2x4 caenorhabditis elegans. y113g7b.9 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)xFx(H)(Y)Wxx(Q) QCFFHYWLIO	RSIIF
1	GCAIN			
1	Q9N8U9	ck: 7148 len: 301	i Q9n8u9 trypanosoma brucei. possible h-sc (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(D)(H)Wxx(L) ELFODHWTL	YEGFS
1	YTDDK			
1	Q9N8E6	ck: 5512 len: 136	i Q9n8e6 trypanosoma brucei. hypothetical (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(E)(F)Wxx(L) LLFHFWFL	CLFVS
1	KNLFT			
1	Q942I6	ck: 5664 len: 589	i Q942i6 caenorhabditis elegans. hypotheti (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)xFx(D)(F)Wxx(F) ENFYDFWFNF	QSWRE
1	TREDV			
1	O462I8	ck: 7893 len: 219	i O462i8 culicoides varileppensis. glutathio (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (O)xFx(D)(Y)Wxx(Q) ORFADYWIPO	LFAKO
1	MGITLY			

Accession	Protein	Species	Length	Sequence
Q17999	ck: 7776	len: 814	1	Q17999 caenorhabditis elegans. c15a7.1 prot
106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q18843	ck: 7424	len: 660	1	Q18843 caenorhabditis elegans. c54g10.3 prot
9: LALDF	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q9XUC9	ck: 6239	len: 170	1	Q9XUC9 caenorhabditis elegans. ml62.2 prote
83: GSIFG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
045282	ck: 676	len: 705	1	045282 caenorhabditis elegans. c31a11.5 pro
294: IESIL	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
001616	ck: 9544	len: 533	1	001616 caenorhabditis elegans. hypothetical
91: LKLET	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q04622	ck: 3173	len: 326	1	Q04622 trypanosoma brucei. u2 small
129: RELRD	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q95JP5	ck: 5683	len: 235	1	Q95JP5 macaca fascicularis (crab eating mac
16: SCVPI	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q9TGJ4	ck: 134	len: 161	1	Q9TGJ4 taenia crassiceps. nadh dehydrogenas
147: SYEYF	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q9B8W7	ck: 7748	len: 417	1	Q9B8W7 taenia crassiceps. nadh-ubiquinone c
403: SYEYF	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q94TE9	ck: 2807	len: 116	1	Q94TE9 chlorophthalmus agassizii (shortnose
18: ALSSL	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q94T60	ck: 3118	len: 116	1	Q94T60 trachipterus trachipterus (ribbon fi
18: TLSTV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q94T47	ck: 2171	len: 116	1	Q94T47 zu cristatus (scalloped ribbonfis
18: LLSV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q94S17	ck: 2075	len: 116	1	Q94S17 antigonla capros (deepbody boarfl
18: TLSTI	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q94SD5	ck: 3914	len: 116	1	Q94SD5 dactyloptena peterseni (starry fl
18: LLTIV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q96303	ck: 605	len: 534	1	Q96303 arabidopsis thaliana (mouse-ear c
130: AVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
050040	ck: 992	len: 542	1	050040 arabidopsis thaliana (mouse-ear c
130: GVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q96302	ck: 8746	len: 524	1	Q96302 arabidopsis thaliana (mouse-ear c
130: GVMTT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
048640	ck: 8027	len: 524	1	048640 arabidopsis thaliana (mouse-ear c
130: GVMTT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
048639	ck: 4821	len: 521	1	048639 arabidopsis thaliana (mouse-ear c
130: GVMTT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q9ZQX5	ck: 8548	len: 3,449	1	Q9ZQX5 arabidopsis thaliana (mouse-ear c
70: KKLLE	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			
Q9ZWT3	ck: 9638	len: 516	1	Q9ZWT3 arabidopsis thaliana (mouse-ear c
131: TVMTT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)			

1	Q9SQY1	ck: 9617	len: 558	1	Q9sqy1	arabidopsis thaliana (mouse-ear cress)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(L)XFX(R)(F)WXX(F)	GIGGD
1	535: SAWAM								
								LIFMCFWVLL	LSVSF
1	Q9AU01	ck: 5248	len: 540	1	Q9au01	lupinus albus (white lupine). phosph	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(L)XFX(R)(F)WXX(F)	GIGGD
130:	SVLAT							LCFFRFWLGF	
1	Q9AU00	ck: 478	len: 539	1	Q9au00	lupinus albus (white lupine). phosph	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(L)XFX(R)(F)WXX(F)	GIGGD
130:	GTVST							LCFFRFWLGF	
1	Q9AVR0	ck: 6827	len: 536	1	Q9avr0	sesbania rostrata. phosphate transp	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(L)XFX(R)(F)WXX(F)	GIGGD
129:	GVMAT							LCFFRFWLGF	
1	Q9AVQ9	ck: 3365	len: 540	1	Q9avq9	sesbania rostrata. phosphate transp	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(L)XFX(R)(F)WXX(F)	GIGGD
132:	GVMAT							LCFFRFWLGF	
1	Q96243	ck: 8141	len: 524	1	Q96243	arabidopsis thaliana (mouse-ear cress)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(L)XFX(R)(F)WXX(F)	GIGGD
130:	GVMTT							LCFFRFWLGF	
1	Q96264	ck: 8941	len: 524	1	Q96264	arabidopsis thaliana (mouse-ear cress)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(L)XFX(R)(F)WXX(F)	GIGGD
130:	GVMTT							LCFFRFWLGF	
1	Q24029	ck: 4193	len: 538	1	Q24029	lycopersicon esculentum (tomato). phosph	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(L)XFX(R)(F)WXX(F)	GIGGD
129:	GVMTT							LCFFRFWLGF	
1	Q04660	ck: 5621	len: 938	1	Q04660	arabidopsis thaliana (mouse-ear cress)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(L)XFX(S)(F)WXX(F)	MAVVG
856:	VSEFQ							LGFDSEFWLGF	
								LVAAT	
1	Q02055	ck: 2442	len: 542	1	Q02055	catharanthus roseus (rosy periwinkle)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(L)XFX(R)(F)WXX(F)	GIGGD
130:	AVMAT							LCFFRFWLGF	
1	Q9ST22	ck: 2750	len: 536	1	Q9st22	nicotiana tabacum (common tobacco).	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
								(P)XFX(D)(F)WXX(F)	

470: KAIDR	PLFADFWERE	IDS LK	
022301	ck: 33 len: 537	1 022301 medicago truncatula (barrel medic).	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
129: SVMAT	LCFFRFWLGF	GIGGD	
022302	ck: 2173 len: 533	1 022302 medicago truncatula (barrel medic).	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
129: SVMAT	LCFFRFWLGF	GIGGD	
022548	ck: 4793 len: 538	1 022548 lycopersicon esculentum (tomato). in	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
129: GVMAT	LCFFRFWLGF	GIGGD	
022549	ck: 3242 len: 528	1 022549 lycopersicon esculentum (tomato). in	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
135: GVMAT	LCFFRFWLGF	GIGGD	
09XEL6	ck: 7287 len: 396	1 09xel6 triticum aestivum (wheat). high-affi	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
99: SVMAT	LCFFRFWLGF	GIGGD	
092PF4	ck: 2196 len: 382	1 09zpf4 arabidopsis thaliana (mouse-ear cres	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
112: TVMAT	LCFFRFWLGF	GIGGD	
09SWG9	ck: 8707 len: 540	1 09swg9 solanum tuberosum (potato). inorganic	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
129: SVMAT	LCFFRFWLGF	GIGGD	
09LLS5	ck: 6167 len: 537	1 09lls5 nicotiana tabacum (common tobacco).	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
130: SVMAT	LCFFRFWLGF	GIGGD	
09M5K0	ck: 8753 len: 265	1 09m5k0 oryza sativa (rice). phosphate trans	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
99: GVMAT	LCFFRFWLGF	GIGGD	
09M562	ck: 15 len: 393	1 09m562 oryza sativa (rice). phosphate trans	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
99: GVMAT	LCFFRFWLGF	GIGGD	
09M4X4	ck: 7892 len: 709	1 09m4x4 zea mays (maize). transposase. 6/	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(E)(F)WXX(F)		
434: FIHRK	LCFVEFWLRF	DTALE	
09LKL5	ck: 9448 len: 390	1 09lkl5 oryza sativa (rice). phosphate tr	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
99: GVMAT	LCFFRFWLGF	GIGGD	
09ATE1	ck: 2612 len: 323	1 09ate1 oryza sativa (rice). phosphate tr	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
92: SVMAT	LCFFRFWLGF	GIGGD	
09AXE0	ck: 5023 len: 333	1 09axe0 oryza rufipogon (wild rice). phos	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
92: GVMAT	LCFFRFWLGF	GIGGD	
09M1T0	ck: 3263 len: 535	1 09mlt0 arabidopsis thaliana (mouse-ear c	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
130: TVMAT	LCFFRFWLGF	GIGGD	
09LZ39	ck: 3770 len: 428	1 09l39 arabidopsis thaliana (mouse-ear c	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(C)(F)WXX(F)		
182: FFLCY	LSFLCFWLGF	CWNHK	
09ASF0	ck: 5843 len: 794	1 09asf0 oryza sativa (rice). hypothetical	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(Q)		
263: KATAL	LVFERFWHKQ	VNQFT	
09ARI9	ck: 1485 len: 543	1 09ari9 lupinus albus (white lupine). pho	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
130: SVLAT	LCFFRFWLGF	GIGGD	
040115	ck: 3380 len: 214	1 040115 luffa cylindrica (smooth loofah)	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(Q)XFX(S)(F)WXX(Q)		
24: PMVKS	QTFDSFWMVQ	HMPPA	
004381	ck: 4865 len: 521	1 004381 arabidopsis thaliana (mouse-ear c	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		
	(L)XFX(R)(F)WXX(F)		
130: GVMAT	LCFFRFWLGF	GIGGD	
043650	ck: 8926 len: 540	1 043650 solanum tuberosum (potato). inorg	
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		

129: SVMTT	(L)xFx(R)(F)Wxx(F) LCFFRFWLGF	GIGGD	Q93WQ8	ck: 8781	len: 525	1	Q93WQ8	triticum aestivum (wheat). Inorga
Q41479	ck: 2263	len: 527	1	Q41479	solanum tuberosum (potato). Inorga			
	(P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGF	GIGGD	130: GVMGT			1	130: GVMGT	
082003	ck: 5788	len: 538	1	082003	lycopersicon esculentum (tomato). pu			
	(P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGF	GIGGD	116: GLNVP			1	116: GLNVP	
129: GVMTT			Q9CZ13	ck: 6690	len: 139	1	Q9CZ13	mus musculus (mouse). 2310040a02r
Q94DB8	ck: 3975	len: 555	1	Q94db8	oryza sativa (rice). putative phosph			
	(P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGF	GIGGD	13: LSYAG			1	13: LSYAG	
136: AVIGT			Q63363	ck: 7377	len: 475	1	Q63363	rattus norvegicus (rat). rat noer
Q949N0	ck: 1025	len: 549	1	Q949n0	arabidopsis thaliana (mouse-ear cres			
	(P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(C)(F)Wxx(L) LIFMCFWVLL	LSVSF	76: IQQRP			1	76: IQQRP	
526: SAMAM			Q9DCJ5	ck: 5548	len: 172	1	Q9DCJ5	mus musculus (mouse). 0610033103r
Q946Y1	ck: 9530	len: 520	1	Q946y1	gossypium arboreum. ceramide glucosy			
	(P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(Y)Wxx(L) LSFSRYWNYL	RKQTF	80: KSHCA			1	80: KSHCA	
306: PLASD			Q91YD4	ck: 1512	len: 1,507	1	Q91YD4	mus musculus (mouse). transient r
Q945E6	ck: 7000	len: 538	1	Q945e6	oryza sativa subsp. indica. phosphat			
	(P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGF	GIGGD	860: MKMAS			1	860: MKMAS	
129: GVMAT			Q56074	ck: 8233	len: 3,005	1	Q56074	hepatitis gb virus a. genome poly
Q944R4	ck: 3251	len: 519	1	Q944r4	arabidopsis thaliana (mouse-ear cres			
	(P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(Y)Wxx(L) LSFGRYWNYL	RKQTF	789: GRLRK			1	789: GRLRK	
305: PLASD			Q9DW56	ck: 4069	len: 655	1	Q9DW56	rat cytomegalovirus (strain maast
Q943Y7	ck: 3376	len: 125	1	Q943y7	thinopyrum intermedium. inorganic ph			
	(P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGF	GVGGD	210: GRALT			1	210: GRALT	
93: GVMGT			Q83896	ck: 5502	len: 595	1	Q83896	ovine adenovirus type 7. terminal
Q93WR0	ck: 9360	len: 525	1	Q93wr0	triticum aestivum (wheat). Inorga			
	(P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGF	GVGGD	241: ECILD			1	241: ECILD	
130: GVMGT			Q91GE6	ck: 8973	len: 243	1	Q91ge6	epiphyas postvittana nucleopolyhe
Q93WQ9	ck: 9330	len: 525	1	Q93wq9	triticum aestivum (wheat). Inorga			
	(P,L,E,C,Q)xFx(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)xFx(R)(F)Wxx(F) LCFFRFWLGF	GVGGD	51: TSFVE			1	51: TSFVE	
130: GVMGT			Q9W7A2	ck: 3716	len: 380	1	Q9w7a2	coturnix coturnix japonica (japan

1			(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (C)XFX(S)(F)WXX(L) CCFSSFWRG	VASSV	
319:	CVGYC				
1			Q90ZD4 ck: 3654 len: 491 i Q90zda gallus gallus (chicken). interferon (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(E)(Y)WXX(Q) PQFCXYWEQ	VQRCG	
446:	LVKLV				
1			O25915 ck: 8156 len: 417 i O25915 helicobacter pylori (campylobacter p (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (O)XFX(S)(H)WXX(F) QDFKSHWRHF	IDSAN	
108:	SUTRK				
1			P72685 ck: 1310 len: 290 i P72685 synechocystis sp. (strain pcc 6803) . (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(H)WXX(Q) LIFWSHWCYQ	TPLSL	
68:	IFLVL				
1			Q9PAP6 ck: 144 len: 544 i Q9pap6 xylella fastidiosa. hypothetical pro (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(H)WXX(F) LRFSSHWEVF	LLSVA	
14:	VDSMN				
1			Q9KVA4 ck: 1939 len: 463 i Q9kva4 vibrio cholerae. rfbg protein. 12/20 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(F)WXX(F) EKFVSFWDPF	VFEYN	
161:	HLLGG				
1			Q9K057 ck: 2237 len: 118 i Q9kq57 vibrio cholerae. hypothetical protei (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(Y)WXX(Q) LGFISYWLWQ	NS	
107:	LALSW				
1			Q9KE45 ck: 7616 len: 155 i Q9ke45 bacillus halodurans. bh1013 protein. (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(H)WXX(L) ERFEDHWVIL	SAKHG	
43:	CAMYA				
1			Q9KB88 ck: 9494 len: 818 i Q9kb88 bacillus halodurans. bh2040 protein. (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(S)(Y)WXX(Q) QLEMSYWAQ	SIANY	
235:	GPIKR				
1			Q9K650 ck: 5308 len: 154 i Q9k650 bacillus halodurans. putative rna po (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(L) LVFKHFWLGL	TSQEI	
110:	ASLSD				
1			Q9UY55 ck: 3936 len: 167 i Q9jy55 neisseria meningitidis (serogroup b) (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(C)(Y)WXX(L) PYFACYWFG	CYPVS	
140:	EGGWM				

1			Q9JVL8 ck: 1872 len: 173 i Q9jvl8 neisseria meningitidis (serogroup (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(C)(Y)WXX(L) PYFICYWFG	FYPAN	
146:	ERGWM				
1			Q9J737 ck: 492 len: 172 i Q9jt37 neisseria meningitidis (serogroup (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(C)(Y)WXX(L) PYFACYWFG	CYPVS	
145:	EGGWM				
1			P72121 ck: 9154 len: 723 i P72121 pseudomonas aeruginosa. outer mem (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(Y)WXX(F) ERFPDYWELF	SPKRG	
491:	GLGHA				
1			Q9I6K4 ck: 1486 len: 392 i Q9i6k4 pseudomonas aeruginosa. probable (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (O)XFX(R)(F)WXX(L) QHFRFWLWL	TTGMN	
49:	LHPAL				
1			Q9I5J3 ck: 7512 len: 358 i Q9i5j3 pseudomonas aeruginosa. hypotheti (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LLEFYRFWGLF	PTVAQ	
89:	AASVF				
1			Q9CNB6 ck: 8093 len: 1,230 i Q9cnb6 pasteurella multocida. rech. 12/2 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(H)WXX(L) ETRVSHWNCL	QYVLT	
856:	VMALP				
1			Q9AAB3 ck: 7467 len: 208 i Q9aab3 caulobacter crescentus. hypotheti (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(F)WXX(L) EGFEDFWGAL	QGETD	
58:	VNGCG				
1			Q98JT9 ck: 592 len: 153 i Q98jt9 rhizobium loti (mesorhizobium lot (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LMFSRYWRML	ASRAA	
113:	HCSSR				
1			Q97KS8 ck: 8409 len: 260 i Q97ks8 clostridium acetobutylicum. abc-t (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(H)WXX(F) EKFSHWHGF	SLTWY	
33:	FSFNS				
1			Q97DB6 ck: 4438 len: 577 i Q97db6 clostridium acetobutylicum. misma (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(Y)WXX(F) ERFKRYWGEF	QKRKR	
64:	HKLNY				
1			Q99VP6 ck: 6701 len: 501 i Q99vp6 staphylococcus aureus (strain n31		

```

1      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
29:  RGLGV      LFFVEFWERF      SYYGM

Q92AV1  ck: 1480  len: 289   1 Q92av1 listeria innocua. lin1818 protein. 1
1      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
175:  HLNFS      EDFLSFWNNL      RMTVT

Q929W8  ck: 3736  len: 391   1 Q929w8 listeria innocua. lin2155 protein. 1
1      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
240:  YRGPf      LSFEDYWPfLL      KFRfLI

Q928I2  ck: 2291  len: 510   1 Q928i2 listeria innocua. pgm protein. 12/20
1      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
31:  AQANK      PNFDRYWANF      PHGEL

Q92937  ck: 7306  len: 315   1 Q92937 archaeoglobus fulgidus. hypothetical
1      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
53:  AVYIL      LFFWSFWNLL      SNLLF

Q92916  ck: 4349  len: 396   1 Q92916 archaeoglobus fulgidus. pheromone sh
1      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
116:  DRDIG      LFTFRWQKL      TFVEK

Q9YFF0  ck: 4613  len: 110   1 Q9yff0 aeropyrum pernix. 110aa long hypothe
1      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
96:  SDVEG      EAFKSEWGAQ      RPRTD

Q9HIS0  ck: 3314  len: 150   1 Q9his0 thermoplasma acidophilum. hypothetical
1      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
26:  LGLRR      LLFPRFWLWF      WLLDH

```

Databases searched:

SWISS-PROT, Release 40.1, Released on 11Feb2002, Formatted on 19Feb2002
 SPTREMBL, Release 19.0, Released on 14Dec2001, Formatted on 19Feb2002

Total finds: 174
 Total length: 211,714,479
 Total sequences: 667,446
 CPU time: 23:58.58

This Page Blank (uspio)

```
1 FINDPATTERNS on plr:* allowing 0 mismatches
1 1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) May
G71532 ck: 4900 len: 934 1 probable threonine/tyrosine-specific protei
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(E)XFX(S)(F)Wxx(F) SLFLP
664: RSSKM ELFLSFWTGF
W5WLR1 ck: 8739 len: 157 1 E5 protein - rhesus papillomavirus (type 1)
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(P)XFX(C)(F)Wxx(F) CCCFC
37: YACKP PPCCCFWLGF
C69762 ck: 9486 len: 492 1 dl-tripeptide ABC transporter (membrane pr)
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(L)XFX(E)(F)Wxx(F) SYGGM
28: RGLFT LFFTEFWERF
C69291 ck: 4349 len: 396 1 pheromone shutdown protein (trab) homolog -
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(L)XFX(R)(F)Wxx(L) TFVEK
116: DRDIG LTFTRFWQKL
G86502 ck: 36 len: 918 1 S/T protein kinase [imported] - Chlamydomon
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(E)XFX(S)(Y)Wxx(F) IPHLN
648: RSSKM ELFLSYWSGF
C72120 ck: 36 len: 918 1 s/t protein kinase - Chlamydomonila pneumoni
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(E)XFX(S)(Y)Wxx(F) IPHLN
648: RSSKM ELFLSYWSGF
AF1811 ck: 1483 len: 356 1 dTDP-glucose 4-6-dehydratase [imported] - A
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(E)XFX(H)(H)Wxx(Q) GKPAN
125: TFTLL EAFRHHWTKQ
AH1381 ck: 2157 len: 510 1 phosphoglycerate mutase homolog pgm [import
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(P)XFX(R)(Y)Wxx(F) PHGEL
31: AOANK PNEDRYWADF
A11750 ck: 2291 len: 510 1 phosphoglycerate mutase homolog pgm [import
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(P)XFX(R)(Y)Wxx(F) PHGEL
31: AOANK PNEDRYWANF
A27671 ck: 7377 len: 475 1 spectrin alpha chain, nonerythroid - rat (f
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(E)XFX(E)(H)Wxx(L) EQFVEHWKEL
76: IQGRP KQLAA
```

```
1 T32253 ck: 8670 len: 491 1 hypothetical protein T15B7.16 - Caenorha
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(L)XFX(S)(F)Wxx(L) KALPA
275: YSSVF LSFVSEFWIDL
S40462 ck: 4989 len: 535 1 t-complex-type molecular chaperone tcp1
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(L)XFX(E)(Y)Wxx(F) IILRE
3: MA LDFDEYWRPF
PC2306 ck: 7087 len: 189 1 dnaJ protein - Synechococcus sp. (strain
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(Q)XFX(R)(Y)Wxx(Q) GAASV
68: RRAEY QRFSTRWQQQ
JC5550 ck: 2256 len: 287 1 dnaJ protein - Synechococcus sp.
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(Q)XFX(R)(Y)Wxx(Q) GAASV
68: RRAEY QRFSTRWQQQ
JC5648 ck: 4120 len: 641 1 terminal protein precursor - Ovine adeno
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(L)XFX(E)(F)Wxx(L) FIDEF
287: ECILD LPFSEFWIEL
T08841 ck: 8233 len: 3,005 1 polypeptide - douroucouli hepatitis GB v
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(L)XFX(S)(Y)Wxx(F) HQAVR
789: GRLRK LEFYSYWCRF
H89844 ck: 6701 len: 501 1 hypothetical protein SA0682 [imported] -
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(L)XFX(E)(F)Wxx(F) SYGGM
29: RGLGV LFFVEFWERF
C97003 ck: 8409 len: 260 1 ABC-type spermidine/putrescine transport
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(E)XFX(S)(H)Wxx(F) SLTWY
33: FSEFNS EKFSHWGHF
C84811 ck: 605 len: 534 1 phosphate transporter (ATPT2) [imported]
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(L)XFX(R)(F)Wxx(F) GIGGD
130: AVMAT LCFFRFWLGF
T07607 ck: 2263 len: 527 1 phosphate transport protein 2 - potato
1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(L)XFX(R)(F)Wxx(F) GIGGD
133: VVMTT LCFFRFWLGF
T07604 ck: 8926 len: 540 1 phosphate transport protein P11 - potato
```

1		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD		91: LKLETT	QTFPDEFWDSQ	LMNNP
129: SVMTT					S56753	ck: 3644 len: 491	1 interferon regulatory factor 3 - chicken
1		T07740 ck: 4193 len: 538	1 probable inorganic phosphate transport prot	1	446: LVKLV		VQRCG
129: GVMTT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD		S73432	ck: 4533 len: 518	1 MG096 homolog D09_orf518 - Mycoplasma pn
1		T01124 ck: 992 len: 542	1 probable phosphate transporter At2g32830 [i	1	251: SLWNF	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(S)(Y)WXX(Q) QEFISYWTGO	GKDVK
130: GVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD		S73923	ck: 7095 len: 417	1 MG288 homolog G07_orf417 - Mycoplasma pn
1		T07892 ck: 33 len: 537	1 probable inorganic phosphate transport prot	1	164: SLTSA	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(E)(Y)WXX(Q) QKFEYWLKQ	GKDVY
129: SVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD		B72719	ck: 4613 len: 110	1 probable bacterioferritin comigratory pr
1		T07894 ck: 2173 len: 533	1 probable inorganic phosphate transport prot	1	96: SDVEG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(F)WXX(Q) EAFKSEFWGAQ	RPRTD
129: SVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD		G70077	ck: 7683 len: 153	1 hypothetical protein yx1k - Bacillus sub
1		T05714 ck: 4793 len: 538	1 probable inorganic phosphate transport prot	1	32: ELEDA	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(F)WXX(L) ESPYRFWAGL	AYDRI
129: GVMTT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD		T29895	ck: 5664 len: 589	1 hypothetical protein F38A5.13 - Caenorha
1		T07164 ck: 5788 len: 538	1 probable inorganic phosphate transport prot	1	222: TREDV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(F)WXX(F) ENFYDEFWNF	QSWRE
129: GVMTT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD		T19595	ck: 676 len: 705	1 hypothetical protein C31A11.5 - Caenorha
1		T05724 ck: 3242 len: 528	1 probable inorganic phosphate transport prot	1	294: IESIL	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PFFGRFWNHL	VMNAF
135: GVMTT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD		B81828	ck: 492 len: 172	1 hypothetical integral membrane protein N
1		T07808 ck: 2442 len: 542	1 probable inorganic phosphate transport prot	1	145: EGGWM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(C)(Y)WXX(L) PYFACYWFGI	CYPVS
130: AVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD		D81049	ck: 3936 len: 167	1 hypothetical protein NMB1733 [imported]
1		T47629 ck: 3263 len: 535	1 phosphate transport protein - Arabidopsis t	1	140: EGGWM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(C)(Y)WXX(L) PYFACYWFGI	CYPVS
130: TVMTT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD		A81923	ck: 1872 len: 173	1 probable membrane protein NMA0784 [impor
1		T34457 ck: 9544 len: 533	1 hypothetical protein T19H12.10 - Caenorhabd	1	146: ERGWM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(C)(Y)WXX(L) PYFICYWFGI	FYPAN
1		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(D)(F)WXX(Q)					

1	528025	ck: 7732	len: 459	1	light harvesting complex II assembly factor	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	ISREN
1	244: KPEKE				PEFGDHWREF		
1	A64690	ck: 8156	len: 417	1	competence locus E - Helicobacter pylori (S	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	IDSAN
1	108: SLTRK				QDFKSHWRHF		
1	S74540	ck: 1310	len: 290	1	hypothetical protein sll0696 - Synechocystis	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	68: IFLVL				LIFFSHWCYQ		TPLSL
1	S73732	ck: 3483	len: 395	1	MG306 homolog A05_corf395 - Mycoplasma pneum	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	169: LEQNH				EGFVSIFYWSEF		AGMIF
1	S46356	ck: 5380	len: 324	1	pupR protein - Pseudomonas putida	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	310: LQLEV				QHFTRYWVTL		KPRMA
1	A11701	ck: 3736	len: 391	1	hypothetical protein lin2155 (import	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	240: YRGPF				LSFEDYWPVL		KFRLL
1	S28472	ck: 8871	len: 460	1	rfbG protein - Vibrio cholerae	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	158: HLLGG				EKFVSEWDPF		VFEYN
1	D82346	ck: 1939	len: 463	1	rfbG protein VC0245 [imported] - Vibrio ch	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	161: HLLGG				EKFVSEWDPF		VFEYN
1	E83776	ck: 7616	len: 155	1	hypothetical protein BH1013 [imported] - Ba	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	43: CAMYA				ERFEDHWVIL		SAKHG
1	S55948	ck: 3922	len: 518	1	hypothetical protein YLR392c - yeast (Sach	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	146: NARLP				PTFNHWOEF		NKIDN
1	F82557	ck: 144	len: 544	1	hypothetical protein XF2449 [imported] - xy	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	14: VDSMN				LRFSHWTFV		LSVA
1	F82112	ck: 2237	len: 118	1	hypothetical protein VC2147 [imported] -	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	107: LALSW				LGFIYSWLMQ		NS
1	C84898	ck: 3197	len: 787	1	hypothetical protein At2g46060 [imported	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	629: LSYNV				LQFMDFWLSF		MAVVG
1	C84582	ck: 6852	len: 550	1	hypothetical protein At2g19880 [imported	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	280: PLASD				LSFGRYWNYL		RKQTF
1	AD0772	ck: 9075	len: 703	1	hypothetical protein STY2348 [imported]	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	672: DCLDH				LHFDSEWKR		REAWY
1	T20228	ck: 7424	len: 660	1	hypothetical protein C54G10.3 - Caenorha	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	9: LALDF				LFRRFRWRL		KILFP
1	H87334	ck: 7467	len: 208	1	conserved hypothetical protein CC0691 (1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	58: VNGCG				EGFEDFWGAL		QGETD
1	E86254	ck: 9544	len: 783	1	hypothetical protein [imported] - Arabid	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	158: NGSDS				LYFGRFWTHL		VMAYA
1	C96782	ck: 7509	len: 654	1	unknown protein F22H5.13 [imported] - Ar	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	470: KAIDR				PLFADFWEF		ISLKL
1	A84888	ck: 8740	len: 1,374	1	hypothetical protein At2g45230 [imported	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	596: KAMRG				LGFDADHWIRL		IMECV
1	A84448	ck: 6617	len: 131	1	hypothetical protein At2g03400 [imported	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	
1	49: FLRKS				PSFRSHWKS		KORNL

1	H83554	ck: 7512	len: 358	1	hypothetical protein PA0736 [imported] - Ps	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	AASVF				PTVAQ		(L)XFX(R)(F)WXX(F)		LEFVDFWLYQ		106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	F83173	ck: 9154	len: 723	1	outer membrane protein OprC PA3790 [importe	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	GLGHA	491: GLGHA			SPKRG		(E)XFX(D)(Y)WXX(F)		EGFKEFWIGL		106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	F83610	ck: 1486	len: 392	1	probable fatty acid desaturase PA0286 [impo	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	LHPAL	49: LHPAL			TTGMN		(O)XFX(R)(F)WXX(L)		QCFEHWLIG		106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	S73657	ck: 3241	len: 260	1	MG288 homolog F04_orf260V - Mycoplasma pneu	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	M	2: M			KKDV		(O)XFX(E)(Y)WXX(Q)		QLEFSYWAQ		106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	E69365	ck: 7306	len: 315	1	hypothetical protein AF0925 - Archaeoglobus	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	AVYIL	53: AVYIL			SNLLF		(L)XFX(S)(F)WXX(L)		LVFKHFWLGL		106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	T01809	ck: 5621	len: 938	1	hypothetical protein A_TM021B04.3 - Arabidoc	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	VSFRO	856: VSFRO			LVAAT		(L)XFX(S)(F)WXX(F)		ERFRYWGEEF		106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	T01083	ck: 8548	len: 3,449	1	hypothetical protein T10P11.5 - Arabidopsis	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	KLKLE	70: KLE			RSSSS		(L)XFX(R)(Y)WXX(F)		ERFPDYWELF		106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	T48284	ck: 3770	len: 428	1	hypothetical protein T22P11.220 - Arabidops	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	FELCY	182: FELCY			CWNHK		(L)XFX(C)(F)WXX(F)		EDFLSFWNLL		106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	T39938	ck: 5030	len: 1,649	1	hypothetical protein SPBC23E6.04c - fission	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	EHRLD	337: EHRLD			IKSRK		(E)XFX(S)(Y)WXX(L)		EDFLSFWNLL		106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	S36335	ck: 3173	len: 326	1	U2 snRNP 40K protein - Trypanosoma brucei	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	RELRD	129: RELRD			EILSI		(L)XFX(S)(H)WXX(L)		QTFTRYWSIL		106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	T19297	ck: 7776	len: 814	1	hypothetical protein C15A7.1 - Caenorhabdit	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	106: FTRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)

AG2271 ck: 4679 len: 59 1 CAB/ELIP/HLIP superfamily [imported] - Anab

1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(Q)XFX(H)(F)WXX(L)

50: EAFSG QGFLHFWGIL

AH1862 ck: 580 len: 40 1 CAB/ELIP/HLIP-related protein [imported] -

1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
(Q)XFX(H)(F)WXX(L)

31: EFFSG QGFLHFWNLL

Databases searched:
NBRF, Release 71.0, Released on 15Feb2002, Formatted on 19Feb2002

Total finds: 85
Total length: 96,089,334
Total sequences: 283,138
CPU time: 10:43.16

This Page Blank (uspio)

11AA_SEQUENCE 1.0
ID 2A5E_MOUSE STANDARD; PRT; 387 AA.
AC Q61151;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Serine/threonine protein phosphatase 2A, 56 kDa regulatory subunit,
DE epsilon isoform (PP2A, B subunit, B' epsilon isoform) (PP2A, B
DE subunit, B56 epsilon isoform) (PP2A, B subunit, PP61 epsilon isoform)
DE (PP2A, B subunit, R5 epsilon isoform) (Fragment).
GN PPP2R5E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryonic fibroblast;
RX MEDLINE=97042488; PubMed=8887688;
RA Okamoto K., Kamibayashi C., Serrano M., Prives C., Mumby M.C.,
RA Beach D.;
RT "p53-dependent association between cyclin G and the B' subunit of
RT protein phosphatase 2A.";
RL Mol. Cell. Biol. 16:6593-6602(1996).
RN [2]
RP IDENTIFICATION OF PROBABLE FRAMESHIFT.
RA Hulo C.;
RL Unpublished observations (JAN-2000).
CC -1- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
CC COMPARTMENT. INTERACTS WITH CYCLIN G (IN VITRO).
CC -1- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATES
CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
CC B (THE R2/B/PR55/B55, R3/B'/PR72/PR130/PR59 AND R5/B'/B56
CC AND CELL SIGNALING MOLECULES.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
CC FAMILY.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT WAS
CC INTRODUCED FROM POSITION 112-135 TO EXTEND THE SIMILARITY WITH THE
CC HUMAN HOMOLOG.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U49728; AAB37234.1; ALT_FRAME.
DR MGD; MGI:1349473; PPP2R5E.
DR InterPro; IPR002554; B56.
DR Pfam; PF01603; B56; 1.
DR Multigene family.
FT NON_TER 1 1
FT NON_TER 387 387
SQ SEQUENCE 387 AA; 45497 MW; 780D5404848A548E CRC64;
2A5E_MOUSE Length: 387 May 30, 2002 09:32 Type: P Check: 8649 ..
1 LKDVPTSEQP ELFLKLLQOC CYICDFMDTL SDLKMEYKR STINELVDYI
51 TISRGLTEQ TYPEVVRMS CNIFRTLPPS DSNEFDPEED EPTLEASWPH
101 LQLYVEFFIR FWSQEFQPS IAKKYIDQKE VLQLELEFDS EDPRERDYLK
151 TVLHRIYGEF LGLRAFIKQ INNIFLRFVY ETEHFNQVAE LLEILGSIIN

201 GFALDLKAEH KQFLVKVLP LHTVRSLSLF HAOLAYCIVQ FLEKPSLTE
251 PVIRGLMKFW PKTCSQKEVM FICELEELD VIEPSQFVKI QEPLEKQIAK
301 CVSSPEHQVA ERALYYWNE YMSLIEENS NVILPIMSS LYRISKEHWN
351 PAIVALVYTV LKAEMEMNST MEDELATYK SDRQREK
11AA_SEQUENCE 1.0
ID BRAC_HALRO STANDARD; PRT; 471 AA.
AC P56158;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Brachyury protein homolog (T protein) (AS-T).
GN T.
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyuridae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gastrula;
RA Yasuo H., Satoh N.;
RT "An ascidian homolog of the mouse Brachyury (T) gene is expressed
RT exclusively in notochord cells at the fate restricted stage.";
RL Dev. Growth Differ. 36:9-18(1995).
CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES
CC REQUIRED FOR MESODERM DIFFERENTIATION (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: DEVELOPING NOTOCHORD.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED IN 64-CELL EMBRYOS. REACHES
CC MAXIMAL LEVELS AT THE GASTRULA STAGE, DECREASES IN NEURULA STAGE
CC AND BARELY DETECTABLE IN LARVAE.
CC -1- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D16441; BAA03910.1; -.
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX_1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
DR Developmental protein; Transcription regulation; DNA-binding;
KW Nuclear protein; Activator.
FT DNA_BIND 24 196 T-BOX.
FT DOMAIN 340 343 POLY-ASN.
FT DOMAIN 348 353 POLY-ASN.
SQ SEQUENCE 471 AA; 52417 MW; 096B8BDC42F205C4 CRC64;
BRAC_HALRO Length: 471 May 30, 2002 09:32 Type: P Check: 462 ..
1 MSITNNMESP SDESVRLTLN DRAIMTKFCS LTNEMIVTKS GRMFPVLKL
51 TASGLEPNSM YSFLDFAPA DSNRWKYVNG EWPVGKPEP HAASCVYVHP
101 DSPNFGSHWM KQPVSNKVK LFNKNGGGO QIMNSLHKY EPRIHVVKVG
151 GEASERTIA TFSPPESOFI AVTAYQNEEV TSLKIKHNPF AKAFIDAKER
201 PDQTFHSLA GIPVSSPOVP SWGRNGSTS SARHFTHCNS YGGESELSYV

```
251 QDTAIPSYTS RNCMRNSYRG NARATPYTIP HKELTQCATS FPEPVPNDGF
301 YPMFPNSELL PRTTLNNYSP AMGAYTNSSI VTSSDIQSGN NNNFFYSNNN
351 NINTTDEVPT TYMTDNFSNF YNQSNSGMP GTTYLPYQSS PVNQFYSYOP
401 PYSTEIADIS PTQODIINAQ NPYQTAWTPP LSYDGCSTMY NSITPYSSSG
451 ESTTSEMTLL ATARYLQNLR L

!!AA_SEQUENCE 1.0
ID CHS1_TUBUN STANDARD; PRT; 189 AA.
AC P55003;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Chitin synthase 1 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
transferase 1) (Fragment).
GN CHS1.
OS Tuber uncinatum (Burgundy truffle).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Pezizomycetes;
OC Pezizales; Tuberales; Tuberales; Tuber.
OX NCBI_TaxID=36049;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95031022; PubMed=7944356;
RA Mehnann B., Brunner I., Braus G.H.;
RT "Nucleotide sequence variation of chitin synthase genes among
RL Appl. Environ. Microbiol. 60:3105-3111(1994).
CC -1- FUNCTION: PLAYS A MAJOR ROLE IN CELL WALL BIOGENESIS.
CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + ((1,4)-(N-acetyl-
CC beta-D-glucosaminyl))(N) = UDP + ((1,4)-(N-acetyl-beta-D-
CC glucosaminyl))(N+1).
CC -1- SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X78101; CAA54994.1; -.
CC InterPro: IPR002923; Chitin_synth.
CC Pfam: PF01644; Chitin_synth; 1.
CC ProDom: PD002998; Chitin_synth; 1.
CC Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
CC Multigene family.
CC FT NON_TER 1 1
CC FT NON_TER 189 189
CC SQ SEQUENCE 189 AA; 21187 MW; 940F6C0ABA7422C9 CRC64;

CHS1_TUBUN Length: 189 May 30, 2002 09:32 Type: P Check: 8541 ..

1 EMNFTRTMHG VMRNIAHFC SRSKRVWGD GWOKIVCTIV ADGRKRVHPR
51 VLDAFPAMGV YQAGIAKNSI NGREVKAHVY EYTTQVSLDS DLKFKGAEKG
101 IVPVQIMFCL KELNAKLINS HRCIFQRFWS TLNPNVCILL DVGTRPGHDS
151 LYHLMKAFDT DSNVGACGE IKAMKGGM GLFNPLVAS

!!AA_SEQUENCE 1.0
ID DNAJ_SYNP7 STANDARD; PRT; 287 AA.
AC P50026;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chaperone protein dnaJ.
```

```
GN DNAJ.
OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97382457; PubMed=9240461;
RA Oguchi K., Nimura K., Yoshikawa H., Takahashi H.;
RT "Sequence and analysis of a dnaJ homologue gene in cyanobacterium
RL Synechococcus sp. PCC7942.";
RL Biochem. Biophys. Res. Commun. 236:461-466(1997).
RN [2]
RN SEQUENCE OF 1-189 FROM N.A.
RX MEDLINE=94271242; PubMed=8003021;
RA Nimura K., Yoshikawa H., Takahashi H.;
RT "Sequence analysis of the third dnaJ homolog gene in Synechococcus
RL sp. PCC7942.";
RL Biochem. Biophys. Res. Commun. 201:848-854(1994).
RN [3]
RN ERRATUM.
RX MEDLINE=95110355; PubMed=7811295;
RA Nimura K., Yoshikawa H., Takahashi H.;
RL Biochem. Biophys. Res. Commun. 205:2016-2017(1994).
CC -1- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DNAJ FAMILY.
CC -1- SIMILARITY: CONTAINS 1 J DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB003519; BAA21679.1; -.
CC EMBL: D29968; BAA06235.1; -.
CC DR HSSP: P25685; 1HDJ.
CC DR InterPro: IPR002939; DnaJ_C.
CC DR InterPro: IPR001623; DnaJ_N.
CC DR Pfam: PF00226; DnaJ; 1.
CC DR Pfam: PF01556; DnaJ_C; 1.
CC DR SMART: SM00271; DnaJ; 1.
CC DR PROSITE: PS00636; DNAJ_1; 1.
CC DR PROSITE: PS50076; DNAJ_2; 1.
CC KW Chaperone; DNA replication; Heat shock.
CC FT DOMAIN 4 73
CC FT DOMAIN 287 287
CC SQ SEQUENCE 287 AA; 31980 MW; 2927A269B2E6C546 CRC64;

DNAJ_SYNP7 Length: 287 May 30, 2002 09:32 Type: P Check: 2256 ..

1 MONFRDYVAL LGIPQADQA AIKAAFRKLA RQCHPDINPG DRQAEERFKQ
51 ISEAYEILSD PDRAEYQRF SRYWQQGA SVGSDDYDGD FPDFDIFVDE
101 LIGRTYERS PRSARRSAA TSSALSRDIE RSLEYDPKTA LQGGSAQLQL
151 EDGRLLFVDI PAGIQAGEYL RLRGGGIRGG DLLLRVQLQA SNFYQVGSVDV
201 IYTLNVSAM AVLGQVTVR TLDGPVQMKL PASLRSGORL RLAGKGYSKP
251 SGDRGDQIVV IQLQLPRLS PEERQLYEQL RSLQQR

!!AA_SEQUENCE 1.0
ID IRF3_CHICK STANDARD; PRT; 491 AA.
AC Q90643;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Interferon regulatory factor 3 (IRF-3).
OS Gallus gallus (Chicken).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=95334365; PubMed=7541908;
 RA Grant C.E., Vasa M.Z., Deeley R.G.;
 RT "cIRF-3, a new member of the interferon regulatory factor (IRF)
 RL family that is rapidly and transiently induced by dsRNA.";
 CC Nucleic Acids Res. 23:2137-2146(1995).
 CC -1- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH
 CC OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT
 CC FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT
 CC (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-1.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U20338; AAA86995.1; -.
 DR HSSP: P15314; 1IF1.
 DR InterPro: IPR001346; IRF.
 DR Pfam: PF00605; IRF; 1.
 DR PRINTS: PR00267; INTERNEGET.
 DR PRODOM: PD002355; IRF; 1.
 DR SMART: SM00348; IRF; 1.
 DR PROSITE: PS00601; IRF; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein; Activator.
 FT DNA_BIND 14 112 TRYPTOPHAN PENTAD REPEAT.
 SQ SEQUENCE 491 AA; 54441 MW; CAE0C2AA8BE976D9 CRC64;
 IRF3_CHICK Length: 491 May 30, 2002 09:32 Type: P Check: 3644 ..
 1 MAALDSEGD AOKLRFQWLL NAVSSGLYRG LQWIDPRRI FRIPMKHNR
 51 KDVTSDEVEI FKAWAKASGR YEGNAEDPAK WKTNRCALR STMFMILLD
 101 RSKCNDDBPK VYAVASGVN DRSGGPAVAG ALQQPQLLL NHHDLALENT
 151 PTDSTGVAA AALTQVDL LQSVLQHCNI SALGSOPTLW AHTGDALPED
 201 ALLPGQDGC LPPQFQDWR QLEPILLGN QPLTGGCGG DGAGALPVSE
 251 ECAIPAPSPA EELLFQSANP APPPAGDIG GLPPLDITI YRGKMYQE
 301 QVDSRCVLA YQPLDPVAE QRLVLPSPA SLDPDQRRY TEDLLEAVGL
 351 RLEQAGQLL ATRLKCKVE WALSQLEGG EPPLNLHRD QETTFDERV
 401 FCTELRFRD SRERSPDFT IFLCFGQCF STKPKSKLI LKLVDPQCE
 451 YWYEQVRGG ASSLSGNVS LQLDSFNLF ELIEQYHMQT D
 11AA_SEQUENCE 1.0
 ID PKN1_CHLMU STANDARD; PRT; 934 AA.
 AC Q9PK92;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable serine/threonine-protein kinase 1 (EC 2.7.1.-).
 GN TC0575.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83560;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Mopn / Nigg;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE002326; AAF73573.1; -.
 DR TIGR: TC0575; -.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding; Complete proteome.
 FT DOMAIN 4 296 PROTEIN KINASE.
 FT NP_BIND 10 18 ATP (BY SIMILARITY).
 FT BINDING 33 33 ATP (BY SIMILARITY).
 FT ACT_SITE 138 138 BY SIMILARITY.
 SQ SEQUENCE 934 AA; 107498 MW; F08AEFD10A47E05C CRC64;
 PKN1_CHLMU Length: 934 May 30, 2002 09:32 Type: P Check: 1219 ..
 1 MORYELIRLI GRGGMGEYLL AHDKACSRV ALKKIRBDLS DNPILKRRFL
 51 REAKIAADLI HPGIVPVYSI CSDGESVYTT MPYIEGSLK HLKSWQKE
 101 ILSKLEERT SVKAFPLIFD KICATVEYIH SKGVLRDLK PDNILLGLFG
 151 EVVIDWGA AIFKAKELQQ EKDEEGFSY GQKNICYSSM TVPKIVGTP
 201 DYMAPESLG AEASEKTDIY ALGLLYQML TLFSPYRRKK GRKLPYEDSI
 251 LSPLEMAPYR EIPPSLQIA MKAIADVQV RFSSVQELRK ALQPHUGES
 301 EMTTRDILST KDRKNWKYYE PILSRFPV LASSPAQWYN FMLSMEVNS
 351 SVRECSVTK SSVQEGVGIF FPPSKKADKG EFGYGLWF SSQNNELSVS
 401 LIKNGIEIQK ESQGIIPQOS RFAISIEKN NKITVYDQI LFIHIDYLP
 451 SLGERIEIIL QDLQISNIT ILESIGALRV SCLAVPDAFL AEKLYDQAR
 501 FYRKIRDSFP GRKESYEAF RLGVTLTLQI EEQGDLMQA LSTFDLLHGS
 551 TGAPLEYLCK ALVYQRNGSF VEEIRSLLLA LKRYPOHPEI PRKDHLCFR
 601 LYDSLHKHRS EALVFMILLI WIAPEKIGLR EERFLEFLH HROOSTLFCR
 651 IDKTPLOFRS SKMELFISFW TGFTLFPEL FQARDLRDY QALIDIFYV
 701 CASGNKEVPS QFAEDLAFV DEVVFKSLH NQGEELVLF VQGLAALQNR
 751 EYRQAKEFIS AVFPALQLYA LDLFSLQAFI DEVVAFSDF LQDIYNSASA
 801 EDHKHVLVYM IQVSLMNQDL KQAYELLSKN FPQDKGLIEY SEAFVLMGCY


```

851 LALTGDRSAV KAHESRCQFK YGRSALIGKC IDDDSLDYLE GLVWEKKKT
901 LFQSTFLLRC LHAPEKRYEV YRQAYISMEN SFFG

11AA_SEQUENCE 1.0
ID PKN1_CHLPN STANDARD; PRT; 932 AA.
AC Q9Z986; Q9K216;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable serine/threonine-protein kinase 1 (EC 2.7.1.-).
GN CPN0095 OR CP0679 OR CPJ0095.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.R., Peterson J., Uitterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AEO01595; AAD18248.1; ALT_INIT.
DR EMBL; AEO02226; AAE73695.1; -.
DR EMBL; AP002545; BAA98305.1; ALT_INIT.
DR TIGR; CP0679; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding; Complete proteome.
FT DOMAIN 4 291 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 138 138 BY SIMILARITY.

```

```

SQ SEQUENCE 932 AA; 107409 MW; 47AAAF87DF6CC5DEF7 CRC64;
PKN1_CHLPN Length: 932 May 30, 2002 09:32 Type: P Check: 8398 ..

1 MERYDIVRII GKGMGEVYL AYDVCSSKV ALKKIREDLA ENPLKRRFL
51 REARIAADLI HPGVVPVYTI YSEKDPVYIT MPYIEGYTLK TLKSWQKE
101 SLSELAECT SVGAFLSIFH KICCTIEYVH SRGILHRDLK PDNILLGFS
151 EAVILDWGAA VACGEEDLL DIDVSKEVL SSRMTIPGRI VGTPDYMAPE
201 RLIGHPASKS TDIYALGVVL YQMLTISFPY RRRKGKKIYL DGQRIPSPQE
251 VAPYREIPEF LSAYVMRLA VDPQERYSSV TELKEDIESH LKGSFKWTLT
301 TALPPKKSSS WKLINEPILLS KYFPMLEVSP ASWYSLAISN IESFSEMRLE
351 YTLSSKGLNE GFGILLPTSE NALGDFYQG YGFWLHIKER TLSVSLVKNs
401 LEIQRCSQDL ESDKETFLIA LRQHNHSLSL FVDGTTWLII MNYLPSRSGR
451 VAIIVRDMED ILEDIGIFES SGLSRVSCLA VPDAFLAEKL YDRALVLYRR
501 IAESEPPGRKE GYEAERFRAGI TVLEKASTDN NQGFALAIIE EFSKLHDGVA
551 APLEYLGKAL VYORLQEYNE EIKSLLALR RYSQHPEIFR LKDHVYRLH
601 ESFYKRDRLA LVEMLIVLEI APQAITPGQE EKILWLKDK SRATLFCLLD
651 PTVLELRSSK MELFLSYWSG FIPHLNSLFH RAMQDSQDVA LIEIFYACD
701 LHKQQLSSC IDIFKESLED QKATEIEVEF SPEDLGAFLE AIQSIENKED
751 AEKIFVSNDQ LSPILLVYIF DLFANRALLE SQGEAIFQAL DLIRSKVPEN
801 FYHDYLRNHE IRAHLWCRNE KALSTIFENY TEKQLKDEQH ELFVLGYCYL
851 ALIIGAFAAK QHFVDCREDR IPPASLARN YNRGLPKDA LSYQERRLLL
901 RQKFLYFHCL GNHDERDLQ TMYHLLTEEF QL

11AA_SEQUENCE 1.0
ID PKN1_CHLTR STANDARD; PRT; 934 AA.
AC Q84303;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable serine/threonine-protein kinase 1 (EC 2.7.1.-).
GN CT301.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
RX MEDLINE=9900809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: AE001302; AAC67894.1; -.
DR HSSP; Q06486; ICKI.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding; Complete proteome.
FT DOMAIN 4 296 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 138 138 BY SIMILARITY.
SQ SEQUENCE 934 AA; 107665 MW; A06EC6FAB025DADB CRC64;

PKNL_CHLTR length: 934 May 30, 2002 09:32 Type: P Check: 4900 ..

1 MORELIRLI GKGMGEVYL AHDKACSRV ALKRIREDLS GNALLRQRL
51 REAKIADLI HPGIVPYSI CSDGEAVYIT MPTIEGFSK SLKSWQKE
101 VLSKELEKT SVKSFLPIFD KICATVEYIH SKGVLHRDLK PDNILLGLFG
151 EVVIIDWGAA IFKHAKELKL EGDDEAAVSF DERNICYSSM TIPGKIYGT
201 DYMAPESLLG VEASEKTDIY ALGLILYQML TLAFPYRRKK GRKLSYRDV
251 LPIEMSPYR EIPPSLSQIA MKAIAINPAD RESSIOELRQ ALQPYLQGD
301 EWTYKATLMA KEKSCWKYYD PILLSRYFPV LASSPAQWYN FMLSEVEISA
351 STRVEYTVTK SAVHEGMGIL FLPSKEAERG EFCYGYGLMF SVQNHELTS
401 LIKNCIEIQK KSGEMISQOS RFAILIEKSD NRITAVEVEQA LFIHLIDYLP
451 SLGNRLGVII QDLQGSNIA ISESIGALRV SCLAVPDAFL SEKLYDQAAI
501 FYRKIRDSFP GRKESYEAF RLGVTLLTQI EEOGGDLTQA LSSFDYLHG
551 AGAPLEYLQK ALVYQRNGSF VEELRCLLFA LKRYSQHPET PRLEDHLCFR
601 LYDSLHKHRS EALVFMILLI WIAPEKISVR EERFLRIIT HKQATLFCQ
651 VDKAPLQFRS SKMELFLSFW TGFSLELPRL FRAGGLRDY QALADIFYVA
701 GVGSGREAFM QFSTALANVS DETTPESLH NOKVAELMF VKGVEALRNK
751 DYQAKAKLLM KTFPTLQLYA LDMFHIOAF DEETESFIDL LQAIYDPASE
801 EERDHILVYI IQTHLMNRDL ERAVKLLNDR FPLDELAEY SEAFILMGY
851 LALTGDRVVV KAHFSRCRYK YGKSALIGKC VDGDIIFYLD NLVWWEKKMT
901 LFGSYFLRC LNESPRRYEK YRQAYLSMEN NFFD

11AA_SEQUENCE 1.0
ID PUCC_RHOSH STANDARD; PRT; 459 AA.
AC Q02443;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE protein pucc.
GN PUC.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX MEDLINE=93086425; PubMed=1453956;
RA Gibson L.C., McGlynn P., Chaudhuri M., Hunter C.N.;

```

```

RT "A putative anaerobic coproporphyrinogen III oxidase in Rhodobacter
RT sphaeroides. II. Analysis of a region of the genome encoding hemf and
RT the puc operon.";
RL Mol. Microbiol. 6:3171-3186(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX MEDLINE=20115911; PubMed=10648776;
RA Choudhary M., Kaplan S.;
RT "DNA sequence analysis of the photosynthesis region of Rhodobacter
RT sphaeroides 2.4.1.";
RL Nucleic Acids Res. 28:862-867(2000).
CC -1- FUNCTION: PUCC IS REQUIRED FOR HIGH-LEVEL TRANSCRIPTION OF THE
CC PUC OPERON.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68796; CAA48701.1; -.
DR EMBL: AF195122; AAF24245.1; -.
DR PIR; S28025; S28025.
KW Antenna complex; Transmembrane.
FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 179 199 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 350 370 POTENTIAL.
FT TRANSMEM 384 404 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
SQ SEQUENCE 459 AA; 48973 MW; CFDD43326C9C07B1 CRC64;

PUCC_RHOSH length: 459 May 30, 2002 09:32 Type: P Check: 7732 ..

1 MSRIAEHLVR IGPRLFPAD AASDPLPLRK LRLSLFOVA VGMAIVLLVG
51 TLNRVAVIEL KVPASVVGIM ISLPLLPAPF RALIGFKSDT HVSALGWRRV
101 PWIYKGLTAL WGGFAIMPFA LIYLGQGYA EGQPFWLGVG SAALAFLMVG
151 GGVHTIQTVG LALATDLAPR EDQPKVGLM YVLLISMIF ASIGFGWLLD
201 PYDAQLIKV ISGVAVAVEF LNMIALMKME PRNRAFTVKP EKEPEFGDHW
251 REFISRENAL HGLIVIGLGT LFGMADVIL EPHYGGEVLSM TVAETTRLTA
301 TFAGGGLVGF WLASWVLGRG FDLPLMAFLG AAAGLPGFFA IMGATEMTNV
351 WVELLGLTVV GEGGLFSHG TLTAATMLAP KEQVGLALGA WGAQATAAG
401 VAIAAGVLR DILQAMPDLS GYGAPYVA VEALEAGFLF LTMIVILPLL
451 RSALAAARRL

11AA_SEQUENCE 1.0
ID TCE2_AVEA STANDARD; PRT; 535 AA.
AC P54411;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE T-complex protein 1, epsilon subunit (TCP-1-epsilon) (CCT-epsilon)
DE (TCP-K36).
OS Avena sativa (Oat).

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaeae; Avena.
OX NCBI_TaxID=4498;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. PEMI; TISSUE=Mesocotyl, and Coleoptile;
RX MEDLINE=94085629; PubMed=7903257;
RA Ehmann B., Krenz M., Mummert E., Schaefer E.;
RT "Two Trp-1-related but highly divergent gene families exist in oat
RL encoding proteins of assumed chaperone function.";
FEBS Lett. 336:313-316(1993).
CC -1- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
CC ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
CC ACTIN AND TUBULIN.
CC -1- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 KDA THAT
CC FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE TCP-1 CHAPERONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75778; CAA53397.1; -.
DR HSSP; P48424; 1A6D.
DR InterPro; IPR002423; TCP1_cpn60.
DR InterPro; IPR002194; TCP_1.
DR Pfam; PF00118; cpn60_TCP1; 1.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00750; TCP1_1; 1.
DR PROSITE; PS00751; TCP1_2; 1.
DR PROSITE; PS00995; TCP1_3; 1.
KW Chaperone; ATP-binding; Multigene family.
SQ SEQUENCE 535 AA; 59034 MW; 02761226F5F17B81 CRC64;
TCE2_AVE5A Length: 535 May 30, 2002 09:32 Type: P Check: 4989 ..
1 MALDDEYWR PEILREQEK KSRLQGLDAQ KANIAAGKSV ARLRTSLGP
51 KGMKMLQSP DGDVTITNDG ATLELMDVD NQIAKLWEL SRSQDYDIGD
101 GTGVVVMAG SLEQAELKLL ERGIPIRVA EGYEMASRIA VDLESISTK
151 YEESATDIEP LVQTCMTTLL SKIVSRCKRA LAEIAVKAVL AVADLERKDY
201 NDLLIKVEGK VGGKLEDTEL VGGIIVDKDM SHPQMPKRIE DAHIALITCP
251 FEPPKPKTKH KVDIDTVEKF QTLRGQEQKY FDEMVKCKD VGATLVICQW
301 GFDDEANHLL MQRELPAVRW VGVLELELIA IATGRIVPR FQELSTKLG
351 KAGLVREKSF GTTKDRLMYI EKCANSKAVT IFIRGNKMM IETKRSIHD
401 ALCVARNLIY NNSIVYGGGS AEISCSIAVE AAADRHGVE QYAIRAFADA
451 LDAIPLALAE NSGLPRIDTL TVVKSQHVKE NNSRCGIDCN DVGTDNMEQ
501 NVEETLIGKQ QOILLATQVY KMILKIDVI TPSEY
11AA SEQUENCE 1.0
ID TRL2_HUMAN STANDARD; PRT; 1503 AA.
AC O94759;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Long transient receptor potential channel 2 (LTPC2) (Transient
GN receptor potential channel 7) (TrpC7).
TRPC7 OR KNP3.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99026133; PubMed=9806837;
RA Nagamine K., Kudoh J., Minoshima S., Kawasaki K., Asakawa S.,
RA Shimizu N.;
RT "Molecular cloning of a novel putative Ca2+ channel protein (TRPC7)
RT highly expressed in brain.";
RL Genomics 54:124-131(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rlesselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -1- FUNCTION: MAY BE A CALCIUM CHANNEL.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.
CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. LTRPC
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB001535; BAA34700.1; -.
DR EMBL; AP001754; BAA95563.1; -.
DR MIM; 603749; -.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR000636; Cation_chan_non_lig.
DR InterPro; IPR000086; NUDIX_hydrolase.
DR InterPro; IPR002153; Trans_recep.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF00293; NUDIX; 1.
DR PRINTS; PR01097; TRANSRECEPTRP.
KW Ionic channel; Transmembrane; Ion transport; Calcium channel.
FT DOMAIN 1 752
FT TRANSMEM 753 773 POTENTIAL. CYTOPLASMIC (POTENTIAL).
FT DOMAIN 774 795 POTENTIAL. EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 796 816 POTENTIAL. POTENTIAL.
FT DOMAIN 817 820 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 821 841 POTENTIAL. POTENTIAL.
FT DOMAIN 842 896 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 897 917 POTENTIAL. POTENTIAL.
FT DOMAIN 918 936 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 937 957 POTENTIAL. POTENTIAL.
FT DOMAIN 958 1025 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1026 1046 POTENTIAL. POTENTIAL.
FT DOMAIN 1047 1503 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 1189 1189 R -> Q (IN REF. 2).
SQ SEQUENCE 1503 AA; 171225 MW; AD329AE79F1A71B5 CRC64;
TRL2_HUMAN Length: 1503 May 30, 2002 09:32 Type: P Check: 5055 ..

1 MEPSALRKAG SEQEGFEGL PRRVTDLGMV SNLRSNSSL FKSRLQCPF
51 GNNDQESLS SWIDENIKKK ECVYFEVSSK LSDAGKVCQ CGYTHEQHLE
101 EATKPHTEGQ TQWDPKKHVQ EMPTDAFGDI VFTGLSQVK KYRVQSQDTP
151 SSVIYHLMTO HMGLDVNLIL ISVTGAKNF NMKPRLSIF RGLYKVAQT
201 TGAWITIGGS HTGVAKQVGE AVPDFSLSSS YKEGELITIG VATWGTVHRR
251 EGLIHPGSGF PAEYILDEDEG QGNLTCLDSN HSHFILLVDG THGOYGEIIP
301 LRTREKEFIS EQTKERGVVA IKIPICVVL EGGEPTLHTI DNATNGTPC
351 VVVEGSGRVA DVIAOVANLP VSDITISLIQ QKLSVEFQEM FETFTESRIY
401 EMTKRIODIV RRROLTVER EKGQGOQDQD VALLOALLKA SRSQDHFGHE
451 NMDHQLKLAIV AMNRVDIARS EIFMDEMOWK PSDLHPTMTA ALISNKPEFV
501 KLFLENGVQL KEFVTWDTLL YLYENLDPSG LFIHSLQKVL VEDPERPACA
551 PAAPRLQMHV VAQVLRLLG DETQPLYP RP RNDRLRLLL PVPHVKLNQ
601 GVSLSLYKR SSGHVTFTMD PIRDLIHWI VQNRRELGI IMAOSQDCIA
651 AALACSKILK ELSKEEDTD SSEMLALAE EYEHRAIGVF TECYRKDEER
701 AQKILTRVSE AMGKTTCLOL ALEAKDMKFV SHGIGQAFLT KYWVGQLSVD
751 NGIMRVTLGM IAFPLLTGL ISFREKRLQD VGTPAARARA FETAPVVVFH
801 LNLSYFAFL CLFAVYLMVD FQPVPSWCEC AIYLMFSLV CEEMROLFYD
851 PDECGLMKKA ALYPSDFWNK LDVGAILLFV AGLCRLIPA TLYPGRVILS
901 LDFIECLRL MHITISKTL GPKIIIVKRM MKDVFFFLFV LAVWVVSFGV
951 AKQALILHNE RRVDMLEFGA VYHSYLTIFG QIPGYIDGVN FNEHCSPNG
1001 TDPYKPKCPE SDATQORPAF PEMLVLLLC LYLLFTNILL LNLILAMFNY
1051 TFOQVOEHND QIMKFORHDL IEEYHGRPA PPPFILLSHL QLEIKRVVLK
1101 TPAKHKQOL NKLEKNEEA LLSWEIYLIKE NYLONROFOQ KORPEOKIED
1151 ISNKVDAMVD LLDLDPLKRS GSMQORLASL EEOVAQTARA LHWIVRTLRA
1201 SGFSSEADVP TLASOKAAEE PDAEPGGRKK TEEPGDSYHV NARHLLYPNC
1251 PVTRPPVPE KVPWTEFLI YDPPFYTAER KDAAMDPMG DTLEPLSTIQ
1301 YNVVDGLRDR RSFHGPYTVQ AGLPINPMGR TGLRGGSLS CGPNHTLYP
1351 MVRMRNRNED GATCRKSIAK MLEVLVVKLP LSEHWALPGG SREPGEMLP
1401 KLRILROEH WPSFENLLKC GMEVYKGYMD DPRNTDNAMI ETVAVSVHFQ
1451 DQNDVELNRL NSNLHACDSG ASIRMOVVDR RPLIYANKHT LIQKAAAEFG
1501 AHY

11AA_SEQUENCE 1.0
ID VES_RHPV1 STANDARD; PRT; 157 AA.
AC P24834;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, last sequence update)
DT 01-OCT-1996 (Rel. 34, last annotation update)
DE Probable E5 protein.
GN
OS Rhesus papillomavirus type 1 (Rhpv 1).
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.

OX NCBI_TaxID=10570;
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=91135018; PubMed=1847267;
RA Ostrow R.S., Labresh K.V., Faras A.J.;
RT "Characterization of the complete Rhpv 1 genomic sequence and an
RL integration locus from a metastatic tumor.";
RL Virology 181:424-429(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60184; AAA79315.1; ALT_SEQ.
DR EMBL; M60184; AAA79316.1; ALT_SEQ.
DR PIR; F38503; W5WLR1.
KW Early protein.
SQ SEQUENCE 157 AA; 17398 MW; AC7AA67158844686 CRC64;

VES_RHPV1 Length: 157 May 30, 2002 09:32 Type: P Check: 8743 ..

1 MVVCIGTQWS HFKPVHTLNS IQVLCKANCC CYACKPPFC CFWLCFCCCF
51 CLALCFVHLL SRCFCVFPVC LSYAAYAVVL GVHSEPVCSF WSVFVLFENP
101 VAFDPACPO CGLOQNDVNT AHRHVIISYF AIYAVNIYFV LALLVGAFAK
151 ATSRART

11AA_SEQUENCE 1.0
ID VP41_BPAPS STANDARD; PRT; 460 AA.
AC O9TI07;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Putative protein P41.
GN 41.
OS Bacteriophage APSE-1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.
OX NCBI_TaxID=106199;
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=99420383; PubMed=10489345;
RA van der Wilk F., Dulleman A.M., Verbeek M., van den Heuvel J.F.J.M.;
RT "Isolation and characterization of APSE-1, a bacteriophage infecting
RT the secondary endosymbiont of acyrthosiphon pisum.";
RL Virology 262:104-113(1999).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF157835; AAF03984.1; -.
DR KW Hypothetical protein.
SQ SEQUENCE 460 AA; 51920 MW; 9572842E1FFDE31 CRC64;

VP41_BPAPS Length: 460 May 30, 2002 09:32 Type: P Check: 4265 ..

1 MKTFPRPYQ HLIINHLDI KRSNWAGMG MGRATATLTA LENLYLSGSE
51 TKPTLVLAFL RVAQSTWPE ALKWSHLRNI EVOPIIGSAK ARIALKNTH
101 ASVFTVNYDN LVLVDILGD TWPEGTIAD ESTRLKSFRL RKGKRTAAL
151 AKIAHKSVMR WVNLTGTSP NGIMDLWGQA WFDQGERLG RTYNAFTSRW

```

201 FKRIQLPGQQ WSRFEPLGFA HLIQIPLALSD VTLSDAADW FDIDEPIHNH
251 INVELPAKAR AHYHAMEKEL FLELGESAIE ALNAAKTIK ILQIASGAIY
301 SDDNRNWTET HDAKIQALES IVNESGTPV LVAYHWKHDL ERLKAFPRK
351 KNLDPARTL TDMNNGKIPL LEAHPASCGH GLNLQDGCNI LVFFSHWMDL
401 EGYQOIERI GPTROAQAGH NRPVFIHHTV AKDTLDEVVM ERRNSKRAIO
451 DLLLEAMKRX

!IAA_SEQUENCE 1.0
ID Y306_MYCPN STANDARD; PRT; 395 AA.
AC P75343;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG306 homolog (A05_orf395).
GN MPN435 OR MP406.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000039; AAB96054.1; -
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 131 151 POTENTIAL.
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 254 274 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 348 368 POTENTIAL.
SQ SEQUENCE 395 AA; 45633 MW; 9E9C1ABD25B89B16 CRC64;

Y306_MYCPN Length: 395 May 30, 2002 09:32 Type: P Check: 3483 ..

1 MSTPKSASFE TRKNILAFSE FIAFLVVSV LVTVEFLDIK TGDVKTIIINT
51 INRTNMPWIL LIVLGIIVTL AwniIIMWV ARRCFHAPW WEWVLFACV
101 QEFQIVTPLS LGQDPFRLYW FIKKGMKKQT AVLLVSTGA FWNLAQALIT
151 WPSFEVLSON YALLEQNHG FVSYWFSG MIEDVVAAL FIFIAVSKRM
201 HVLISGVNQ FRKWKRPYL TKEQIYQRFI DKAEFNKLYG LEIKRLGLTI
251 EKLNLNLIA VVGYSVFAV FAIVKKENAT NNVIDQYSTA DIFNTITAI
301 TASNFIPLPS GEGATQFVMT SFLNAFKSAV GIESQVKQGV FLWRFLSYVI
351 PALLESICFI GWVVQVVEEF KHPKPLPVV NLINHFWMN KKLHN

```

```

!IAA_SEQUENCE 1.0
ID YA48_MYCPN STANDARD; PRT; 518 AA.
AC P75066;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN048 (D09_orf518).
GN MPN048 OR MP106.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: BELONGS TO THE MG032 / MG096 / MG288 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000013; AAB95754.1; -
DR InterPro; IPR004306; DUF237.
DR InterPro; IPR004319; DUF240.
DR Pfam; PF03072; DUF237; 1.
DR Pfam; PF03086; DUF240; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 518 AA; 58873 MW; A0D753FC77DC22EC CRC64;

YA48_MYCPN Length: 518 May 30, 2002 09:32 Type: P Check: 4533 ..

1 MNEIQTQVNE ANNIFPVEAF KVPKVPKELF GFVNQGFEPK LNPGLNLAD
51 NVASLFEQYS LKOASLKDFD ILLEKKNDIV LEHKVRYNFA LQHFERTYV
101 GTGGEINLQF ALQASTTNFS SLEELQASFS KTGDNLFQAL FWKPTVTKLV
151 SGENDLTHIA QTAIGESLFD SRVDLSASII NSEATLKTAE ATFTQVNLN
201 FKAEREKALA IKKAEERIK KELEEQKRRQ EELSKQQRDK EALQKSLMNF
251 QEFISYWTGQ GKDVKQKEQF IQALEAAFSY MNNEVFNLII AGFRSAIQTY
301 YKDGKADQSQ NAKIAFGEGK IQFPKSGPGL DGIEMSDFLR GNLGTGNAHFD
351 LKLKKVEYKN TQKDAQGNQ KKASINWQAK QNNFPFRQYN PWDFSFEVEL
401 KYEGSYGLYP GARFLNLFSS LGIPNDWKE MSVKFVLGK TPQWIDAKPD
451 YPGSLFKFEK NQLKFTPHVK EHVHVENKQF MEKLKSNLH NLELATGATK
501 PPVVDLASYL HYLILNHK

!IAA_SEQUENCE 1.0
ID YC03_KLEPN STANDARD; PRT; 504 AA.
AC Q48449;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 55.8 kDa protein in CPS region (ORF3).
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

```



```

OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHEDID;
RX MEDLINE=95204345; PubMed=7896702;
RA Arakawa Y., Wacharotayankun R., Nagatsuka T., Ito H., Kato N.,
RA Ohta M.;
RT "Genomic organization of the Klebsiella pneumoniae cps region
RT responsible for serotype K2 capsular polysaccharide synthesis in the
RT virulent strain Chedid.";
RL J. Bacteriol. 177:1788-1796(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D21242; BAA04774.1; -.
DR Hypothetical protein.
KM SEQUENCE 504 AA; 55782 MW; AD887595CFDFEFAB CRC64;
SQ
YC03_KLEPN Length: 504 May 30, 2002 09:32 Type: P Check: 2748 ..

1 MPRALSLVF PFSFLRYDSG AENGSCMTIK IARIAVTLGL LSLGAQAYA
51 AGLVNDNDL RNDLAWLSDR GVIHLSTW PLSOEISRA LKKAKPSYSS
101 EQVVLARINQ RLSALKADFR VTGYTSTDKP GTPQGFQGTQ PADNSLGIAF
151 NNGSEWMDVH LQGNVEGGER ISNGSRFNAN GAYGAVKEFN QWLSFGQVPQ
201 WNGPGYEGSL IRGDAMRPMT GELMQRAEQA APETWMLRWV GPWQYQISAS
251 QMNQYTAVPH AKIIGRFTF TPFQSLLEGA SRIMQWGEGC RPOSFSSFMD
301 GFTGHDNTGT DNEPGNQLAG FDEKFKLEPT LGWPVSFYGQ MVEDESGYL
351 PSANMFLGGI EGHGHWGKDA VMWYVEAHDY RTNMSRTNYS YTHHIYKDG
401 YQGGYPLGDA MGGDGLFAG KVELITENNQ RWSTRLAYAK VNPQDSINK
451 AFPHSDTLKG VQLGWSGDVY QSVRLNTSLW YTNANNSDSD DVGASAGIEI
501 PPSL

11AA_SEQUENCE 1.0
ID YC34_MYCPN STANDARD; PRT; 417 AA.
AC P75537;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN234 (G07_orf417).
GN MPN234 OR MP597.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: BELONGS TO THE MG032 / MG096 / MG288 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000058; AAB96245.1; -.
DR InterPro; IPR004306; DUF237.
DR InterPro; IPR004319; DUF240.
DR Pfam; PF03072; DUF237; 1.
DR Pfam; PF03086; DUF240; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 417 AA; 47650 MW; 8880670173D369C9 CRC64;
YC34_MYCPN Length: 417 May 30, 2002 09:32 Type: P Check: 7095 ..

1 MQFHAVYQNN DTKANLDFAL NISTINFATL QELQNSFDLQ GSDLTAGLFF
51 KYSVNKLTSQ TNDLTTIAKT ALGENITQKQ VSLTQSIIRP RLEAAKTQYK
101 QDIAPFAKE RQALAQHLK EIEAKQRAE QLLKEQGEAE KRROEVKNV
151 AETQGFNDSL TSAQKFEYV LKQGDVYTK VELIQALKSS FERNQNRFFN
201 FLIAGERTAI DWYVNOEKNV TTAKNNAFGK NGIOFPVAGF QGIYMSQWLR
251 DELSGKTDIR LNLKLSYQV ENKNSSINWN KOKRIEIKQV KPFNYSFEIN
301 LKRTGSYNVS LMYLIGAIG GIPTSWGTW DMKFIYDGL DSGIVTKQDY
351 PGSKFEETED KLMFTLHVQK QIKVKEQGFV NLLKGQSLDN LDLRTGTTKP
401 PVYDLASYLH FVILTAK

11AA_SEQUENCE 1.0
ID YCF2_MESVI STANDARD; PRT; 890 AA.
AC Q9MUP8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 103.9 kDa protein ycf2 (RF2).
GN YCF2.
OS Mesostigma viride.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
OC Mesostigmatales; Mesostigmataceae; Mesostigma.
OX NCBI_TaxID=41882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIES-296;
RX MEDLINE=20150907; PubMed=10688199;
RA Lemieux C., Otis C., Turmel M.;
RT "Ancestral chloroplast genome in Mesostigma viride reveals an early
RT branch of green plant evolution.";
RL Nature 403:649-652(2000).
CC -1- SIMILARITY: BELONGS TO THE YCF2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF16114; AAF43852.1; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003960; AAA_sub.
DR InterPro; IPR003959; AAA_subfam.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00674; AAA; 1.

```

KW Chloroplast; Hypothetical protein.
 SQ SEQUENCE 890 AA; 103935 MW; 43CAEE991AF2C4B CRC64;

YCF2_MESVI Length: 890 May 30, 2002 09:32 Type: P Check: 5630 ..

1 MOQNMSLSN MIKEKEVEF HVPALLIPTR YREIYREIDR APKLEQSFY
 51 TIFIKWNIT EKVTVTDVFY YSSVEDPLGL QSPRHSRWAQ NTDEQRKQA
 101 RIISQSVYKP NSFIRITSNK FTRVLITSLF VLTFFQSNV FQSPYFESQ
 151 INGELFKKMN NINQLEINST NTCIIKITDL LTHQSKELL GNSLPTNSIL
 201 RWKMHKSISD IKTIENMPEK WIYLPGLGL ENSEQIFSYR NEDICNLTKN
 251 SSNMKIKNN RTILDPKIDK YPLVLKNKEN ILFFHKLPID FFLFPYIVLR
 301 IWLAPVFLW WSYQFNSEK ENIKNNLKNL HDIEISTIQ FVAKAIFRD
 351 ICGMESLKQE LATVAFLLKQ KNYSNSYPMG YLFAGPPGTG KTIMAKAMSY
 401 EAETPYLYVE GSGFRCREEG VANARVDLDF KQIQNISPCI LYIDEIDSTA
 451 ERREANKQL EQLKTIGDSI EGSNINIDOK PSDTVLMQFL IYMDGKKRN
 501 DLIIIGATNR IETLDDAIMR PGRFDROIYF SPPEEERKD ILRIFLRNTK
 551 ALVDDTTKTM MAERSIGLNG CDLRLLADNI LLLSALESRN QOKTIPVINE
 601 DTFDRALERV SRIRHIISNY ELAFGKYDFY RTAYHEAGIA LIHTLPECR
 651 PYYSVKLFPP PLNDRYLEIE RENLKVPSDD IISTNNIDYF VQKIVGLLAG
 701 RAAESILFDF YPGQISTYLN KTYDPNIOGA YNIAHHIVEF GLDVSIVGI
 751 HYLSENENNN IKDPFVTKIN DLIQNKVTLK TNRELKYSQ AASILDENE
 801 WYEQDYPWQF DFIKKQYIYS NESSRLDME IVSILHTLQF YTYDFLKSNE
 851 QLLDHLASLV LKNKSISOKE IHLVNSYGI KIPTKTWKAW

11AA_SEQUENCE 1.0
 ID YCLF_BACSV STANDARD; PRT; 492 AA.
 AC P94408;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical transporter ycf.

OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124189; PubMed=8969502;

RA Yamane K., Kumano M., Kurita K.;
 RT "The 25 degrees-36 degrees region of the Bacillus subtilis chromosome:
 RT determination of the sequence of a 146 kb segment and identification
 RT of 113 genes.";
 RL Microbiology 142:3047-3056(1996).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: BELONGS TO THE PTR2 FAMILY OF TRANSPORTERS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D50453; BAA09000.1; -
 DR EMBL; 299106; CAB12175.1; -
 DR Subtilist; BG12027; YCLF.
 DR InterPro; IPR000109; PTR2.
 DR Pfam; PF00854; PTR2; 1.
 DR PROSITE; PS01022; PTR2_1; 1.
 DR PROSITE; PS01023; PTR2_2; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 67 87 POTENTIAL.
 FT TRANSMEM 88 108 POTENTIAL.
 FT TRANSMEM 110 130 POTENTIAL.
 FT TRANSMEM 160 180 POTENTIAL.
 FT TRANSMEM 185 205 POTENTIAL.
 FT TRANSMEM 232 252 POTENTIAL.
 FT TRANSMEM 255 275 POTENTIAL.
 FT TRANSMEM 294 314 POTENTIAL.
 FT TRANSMEM 333 353 POTENTIAL.
 FT TRANSMEM 367 387 POTENTIAL.
 FT TRANSMEM 392 412 POTENTIAL.
 FT TRANSMEM 434 454 POTENTIAL.
 FT TRANSMEM 464 484 POTENTIAL.
 SQ SEQUENCE 492 AA; 53275 MW; 5042A6590D30E033 CRC64;

YCLF_BACSV Length: 492 May 30, 2002 09:32 Type: P Check: 9486 ..

1 MASIDNESII KSPQKGFEG HPRGLFTLF TEFWERFSY GMRAILLYL
 51 YTEVTNGGLG FDKGTAVAIM SIYGLVYMS TIIGWLADR VEGTANTVFY
 101 GGIFIMFGHI ALAYPGSSIA FYISMVLIIV GTGLKPNVS SVGDLYTKE
 151 DPRDSGFSI FYMGINLGL LAPLIVGTLG QKYNHILGFG AAAGVMLGL
 201 IVEPLTRKKN LGLAGSNVPP PLSKSAIGT GIGVIYAIA VIISVGTGL
 251 TIKRFIDLV ILGILIPVIY FIIMFTSKA DKTEKSLAA YVPLFGAVM
 301 FMAIQEQGAT ILAVYADERI RLSLGGFELQ SSWFQSLNPL FVIFAPIFA
 351 WIMMKLGKRQ PSTPVKFSIG IILAGLSFIT MFPAMQCKE ALVSPMLVL
 401 SFLVLVLGEL CLSPVGLSVT TKLAPAARSA QTMSMFLTN AAAQAINAQV
 451 AGLFDKIPET MYFTIGLIS IVLGGILLL SPVIRAMKG VL

11AA_SEQUENCE 1.0
 ID YF11_MYCPN STANDARD; PRT; 260 AA.
 AC P75275;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MPN511 (F04_orf260V).

OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
 RA Herrmann R.;

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae.";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SIMILARITY: BELONGS TO THE MG032 / MG096 / MG288 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000031; AAB95979.1; -.
DR InterPro; IPR004306; DUF237.
DR Pfam; PF03072; DUF237; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 260 AA; 30422 MW; 85B7FD19D9555FB6E CRC64;

YF11_MYCPN length: 260 May 30, 2002 09:32 Type: P Check: 3241 ..

1 MQAFKEYWQK QKKDVTDKQ LLEALKLSFA KEQNKTFAPL IKNFQDGISN
51 YYPNDQEDQS EAAKTAFTQ GIAFPQSLK GIFMSEWLK QIGEKAKINL
101 DIKSLKVTDS KISPTIKWNK DIGIKRNQDK PYNFRFEIDI EYQGNKLSW
151 LEAIIAKESG IPGEWKGKLN LKFIVDGLS WEIVQKPDYP GSLPQFDDQK
201 QQLLFKLHW EKITVQPEEF MELIKSQNLH NLELRTESTK PPVVDLASYL
251 HYQLKLNQO

!!AA_SEQUENCE 1.0
ID YG44_SCHPO STANDARD; PRT; 1649 AA.
AC 060179;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 186.4 kDa protein C23E6.04c in chromosome II.
GN SPBC23E6.04C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Pohl T.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE BAP28 FAMILY.
CC -1- SIMILARITY: CONTAINS 1 HEAT REPEAT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL023287; CAI18872.1; -.
DR InterPro; IPR000357; HEAT_repeat.
DR PROSITE; PS50077; HEAT_REPEAT; 1.
KW Hypothetical protein.
FT REPEAT 1608 1646 HEAT.
SQ SEQUENCE 1649 AA; 186409 MW; 6B4FD373125216D9 CRC64;

YG44_SCHPO length: 1649 May 30, 2002 09:32 Type: P Check: 5030 ..

1 MASSLQKQLK NIGSNNVLKI NKIRRAPSLI YDPKVAADMD LLEIYVTAVS
51 GFHELAVHEP RLTYFEKTLI GEQSYQVDRV LNRTENEKI DLECVQILRL
101 LAPEFTEKNA LKVLLEMLIR FSIHEYVSDE FILSFLPFHD HPEFARILGC
151 SKPKSRPLLF LENAIIKMPVT LSRADIVHAL SRDKEFFAMF AQEVQNTAES
201 HNMYPELARE WAGTMEVLV AMHSSNEDPN VLDRFELRV SYAVSYVSSI
251 DFOIAGFMLL SSIAASLPLS PSIIPLVSA IFDRLSFDNI KRALICVGHL

301 LQFCSSFEFD HEQLEKLESEF GASSLLIELS QEHRLEDEFV SYWVSLIKSR
351 KQKDKRRLIS LLDTSISQIR VTHEQAKELL SVIPVNDQFK ALQSYRIILD
401 SVIQPERKEG KLDNLINTLQ DKKKSSTFSK KDREVLKKI SEIDSQTSFE
451 QCLAYADSAALDSSVFISL LSKFGDKIPF LLECIANGSE RIILSLIEL
501 RKTIEENKDV DYQIILPVYL YSLQSKDTEV RSRALNLILT FLELRNENLE
551 FSIYGMDDN DNKNLEWLSV VETKYCCSDL LDRSSEIGL DGTLYFSYIP
601 ERLFTEKKPK NASKETAVTS FLSSHACSK LSNVRVLLLE ILTRVHGKVE
651 DAKMQLLPR LEQLSEENSE KFKTVSKREV EALVNCFNHT SFTSLLSFSL
701 SNIVLSQAIC RRIVEIQSHL KDQQLREFVK AVISQDEQPH YYVDVLDSIK
751 IPDTVEKKLI GSVRLVKEKN PALAKRRKID SHIFGDVQR LTRILELLET
801 KNAASYPKLA SPLFEVLNSV IALKEDISS NYLLQLLGL LYEMIGASPI
851 TELSPSIRID TLVGCIRSTN NPQIONKALL LVSAIANAP EAVLHGVMPI
901 FTFMGSTVLS RDDAFSIHVI EQTVKTIVISA LIRLQKDFDS SLVSCFVNA
951 FPHIPQHRRL RLRYLYQTI GSNRLSVYL IQFAEKMLLA KSTNVVAIHD
1001 FCLTLVQSFV VADRIGSINO CSRECKLSLE EQSNSDSNGK AVSLIKLDEL
1051 PMVDLATLG SLRVKYLELI SLVSKAKNFA FDLAKIMENS VDSFEYEQAG
1101 LFESIKLIT LSQSSNEME LGHYVYALRS VITHLPNELE CTYLGKLLHD
1151 ERALLRRKAL SIVQQRVQOG SKVSALTALI PDVYTNISNY SDEETTOLAM
1201 DCLAVMAKRF SASPELFISP IEVSGPYGL KNSARDVQVS AIYCITVLTN
1251 TLAARILPYL ADIVNTSLSI LDDARKDPEG DLELACFSM MIDEFKVLPE
1301 FSSSYVEPTI KCALASDRAF EHDATIGELLE ETIANFIPTL LMKISIFAAW
1351 PECARLGSTA ALRLLELIEL ALQNSRSAL GTVYKSIFKE FLDSFDSRRS
1401 LLEAEDVDNV ETQAVNVFLK FVMKLSDTTF RPLFLHLHSM ALEDLYETDP
1451 SGIVSRQTFE YNFLTIFLDT LKSIYVNYA YVLDDTIELL SSKDTNSEVR
1501 HLVNSSIVSA FENDTEEFWM VPAREGKISP VLIEQIQYAP LLDDKVLVKA
1551 IVELASVASS SDNFRSMNTQ LLQYLRSNI NARLAIQIQ TQLYGRIGEN
1601 WISTLPQSVF FIAELMEDDD DQVETATAEL VRIIDRLGE NESIQDYLT

!!AA_SEQUENCE 1.0
ID YI92_YEAST STANDARD; PRT; 518 AA.
AC P18634;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hypothetical 59.8 kDa protein in SSRI-ATP10 intergenic region.
GN YLR392C OR L8084.13.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favello A., Fulton L., Galtung S., Greco T., Kirsten J.,
RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,

```

RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Rifken L., Riles L., Taich A., Trevaskis E., Vignati D.,
RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-210 FROM N.A.
RX MEDLINE=90277691; PubMed=2141026;
RA Ackerman S.H., Tzagoloff A.;
RT "ATP 10, a yeast nuclear gene required for the assembly of the
RL mitochondrial F1-F0 complex.";
J. Biol. Chem. 265:9952-9959(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U19729; AAB82352.1; -.
DR EMBL; J05463; AAB05631.1; -.
DR PIR; C36445; C36445.
DR SGD; S0004384; YLR392C.
KW Hypothetical protein.
SQ SEQUENCE 518 AA; 59758 MW; 65399EA2FFFC2C23 CRC64;

YL92_YEAST length: 518 May 30, 2002 09:32 Type: P Check: 3922 ..

1 MAPKISISLN PRYNGEIFYSS NDQMSGIVSL QLTALSIK ISVILKGFSE
51 TLTKIDQEYM EQONGMMMPG QDNKSFHTLM KEQRVFPD NVWNLDCSS
101 KPFKVKPGSY NYSFQFDKFP RKPECLKNHT AKTVAFTRS NARLPFTNS
151 HWQEFNKIDN LDLYFYSFGK VIYMOVQLE LKSSSWFKP FHKLIKEIET
201 FEFIEPKDL IIEPDEDNE ELNAPSNSR GNSMTNNEF FNSNLKVPVS
251 KDVKVVNGVG YIKSDRNFQ ANSILIENG IRSRPVSVT STROSTRLVN
301 GMKVPSTYK MGLPDGESNM RIEVRSRLK QYRKDYLFK SGSQNFQYV
351 VVMEGNIASL SKMQITPLK QLNLETTY LSGIANGV SSKLIEIDL
401 NQLKSNKPL DLNEIRENFD GSMFECRL KDHPIRLKY FNEEDYRHG
451 NRLYSFKTCT IKRTFSLQL IEWGINGIRK QSEVNIDPVQ IFCQVREHVE
501 AEALPRVPP PTYTEMAS

11AA_SEQUENCE 1.0
ID YXIK_BACSU STANDARD; PRT; 153 AA.
AC P42302;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yxik.
GN YXIK OR SS8B.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSCL1;
RX MEDLINE=97124196; PubMed=8969509;
RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA Miwa Y., Fujita Y.;
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT containing the lic and cel loci, and creation of a 177 kb contig
RT covering the gnt-sacX region.";
RL Microbiology 142:3113-3123(1996).

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D83026; BAA11690.1; -.
DR EMBL; Z99124; CAB15950.1; -.
DR Subtilist; BG11140; yxik.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 153 AA; 17649 MW; CEC1799C0BF2A4DC CRC64;

YXIK_BACSU length: 153 May 30, 2002 09:32 Type: P Check: 7683 ..

1 METTSSISIG NDMIELKPD QVMACELED AESFYRFWAG LAYDRIMIQV
51 ITTGSFIEDL SEYFEGHAYK VTKLAKREFH FQSILOEADR DIADFLFLLA
101 STNDVFLIT DPQPKSYFS ESKLQCLTDS GERIIMFEYD AVDIYMIGE
151 SYK

11AA_SEQUENCE 1.0
ID Q9X4B6 PRELIMINARY; PRT; 504 AA.
AC Q9X4B6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE K30 CAPSULE BIOSYNTHESIS CLUSTER, PARTIAL SEQUENCE.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E69;
RX MEDLINE=99217011; PubMed=10200954;
RA Drummelsmith J., Whitfield C.;
RT "Gene products required for surface expression of the capsular form of
RT the group 1 K antigen in Escherichia coli (O9a:K30).";
RL Mol. Microbiol. 31:1321-1332(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=E69;
RA Drummelsmith J., Whitfield C.;
RX EMBL; AF104912; AAD21561.1; -.
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
SQ SEQUENCE 504 AA; 55715 MW; 7A1E0A299A1DB841 CRC64;

Q9X4B6 length: 504 May 30, 2002 09:32 Type: P Check: 2369 ..

1 MPRALSLIVF PSFSLRCDG AENGSCMTIK IARIAVTGL LSLGAQAYV
51 AGLVYNDNDL RNDLAWLSDR GVIHLSLSTW PLSQEIIARA LKAKKPSYSS
101 EQVYLARINQ RLSALKADFR VTGYTSTDQP GTPQFGQTQ PADNSLGLAF
151 NNSGEWMDVH LQGNVEGGER ISNGSRFNAN GAYGAVKFWN QWLSFGQYVQ
201 WNGPPEYEGSL IRGDAMPMT GLMQRAEQA APETWMLRWV GPWQYQISAS
251 QMNQYTAVPH AKIIGRTFT TPFQSLLEGA SRIMQWGGEG RPQSFSSFWD
301 GFTGHNTGT DNEPGNQLAG FDFKFKLEPT LGWPVSFYCQ MVGEDESGYL
351 PSANMFLLGV EGHGWGKDT VNMVVEAHD RTNMSRTNYS YTHHIYKGY
401 YQGYPLGDA MGGDGLIAG KVELITEDNQ RWSTRLVYAK VNPEQSLNK
451 AFPHADTLKG LQLGWSGDVY QSVRLNTSLW YTNANNSDSD DVGASAGIEI

```

501 PFSL

```

11AA_SEQUENCE 1.0
ID Q9L1E1 PRELIMINARY; PRT; 70 AA.
AC Q9L1E1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SC3D11.21 PROTEIN.
GN SC3D11.21.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Brown S.P., Harris D.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL157916; CAB76018.1; -.
DR HSSP; P02593; 1CDM.
DR InterPro; IPR002048; EF-hand.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
SQ SEQUENCE 70 AA; 7742 MW; 82A55B9FF1166480 CRC64;

09L1E1 Length: 70 May 30, 2002 09:32 Type: P Check: 7894 ..

1 MADIEARKQ FERITDGDG LITAEFKTA LAQGDWNVV ESVAEAIAG
51 RDLGDKQLS FDEFWAHLNK

11AA_SEQUENCE 1.0
ID P95529 PRELIMINARY; PRT; 492 AA.
AC P95529;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TDNQ PROTEIN (EC 6.3.1.2).
GN TDNQ.
OS Pseudomonas putida.
OG Plasmid pTDNL.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97144524; PubMed=8990291;
RA Fukumori F., Saint C.P.;
RT "Nucleotide sequences and regulatory analysis of genes involved in
RT conversion of aniline to catechol in Pseudomonas putida UCC22
RT (pTDNL).";
RL J. Bacteriol. 179:399-408(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) - ADP + GLUTAMINE +
CC ORTHOPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR EMBL; D85415; BAA12805.1; -.
DR HSSP; P06201; 1LGR.
DR InterPro; IPR001637; GlnA_adenyltn.

```

```

DR InterPro; IPR001691; Gln_synth.
DR Pfam; PF00120; gln-synt; 2.
DR Prodom; PD001057; GlnA_adenyltn; 1.
KW Ligase; Plasmid.
SQ SEQUENCE 492 AA; 54538 MW; 721A977FEAD9CCF CRC64;

```

P95529 Length: 492 May 30, 2002 09:32 Type: P Check: 7071 ..

```

1 MSGKFIKHHG IWSDTQKAAA ADVLNKIEKA GLQWVRLSWP DQYGLRGKM
51 LSVAAALRSAP ASGSEITMAP FFFDTASAIY FNPFSADGGL GSALAGSPN
101 VVMVPDETFE RILPMDRTG WMLADLYMTS GRPFALSPRA ILKRALVEMO
151 DLGYDYQAGL EWEYLTIRIV DPCLPEETLG GPGTPAAPPK VMPYAKGYST
201 LLENHLDVEE PIMAEVROHL LALGMPLRSI EDEWAPRSQME TTFDVMPLGD
251 VADTMVLEFRN AVKQVCRRRG YLASFCKKPA IQGFLASGWH LHOSLTARDS
301 GANAFIPQPG EALSALGRSY VGGLEHACA ASSFTTPTIN GYRRRRPYSL
351 APDRVTVAKD NRAAMARVIS APGDPAARVE NRIGEPANP YLYIASQVFS
401 GIDGIRROLD PGPIQETPYA GDVTILPHNL SEALEVLETS KFFREAFGEE
451 FIRYWMHLRR SEWKREYDAE GQVDFSGDPV TNWEHREYFE LF

```

```

11AA_SEQUENCE 1.0
ID Q44310 PRELIMINARY; PRT; 59 AA.
AC Q44310;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE RNA POLYMERASE SIGMA FACTOR (SIGA).
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PCC 7120;
RX MEDLINE=91193199; PubMed=1901566;
RA Brahamsha B., Haselkorn R.;
RT "Isolation and characterization of the gene encoding the principal
RT sigma factor of the vegetative cell RNA polymerase from the
RT Cyanobacterium Anabaena sp. strain PCC 7120.";
RL J. Bacteriol. 173:2442-2450(1991).
DR EMBL; M60046; AAA22044.1; -.
SQ SEQUENCE 59 AA; 6689 MW; DB46436CD0CE208D CRC64;

```

Q44310 Length: 59 May 30, 2002 09:32 Type: P Check: 4158 ..

```

1 MTDYTKISAS VVEDRNSRWV GFTPQAEIWN GRLAMIGFLA ATLIELFSQ
51 GFLHFWGIL

```

```

11AA_SEQUENCE 1.0
ID Q52209 PRELIMINARY; PRT; 324 AA.
AC Q52209;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PUPR PROTEIN.
GN PUPR.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LOC5358;
RX MEDLINE=94298771; PubMed=8026465;
RA Koster M., Van Klompenburg W., Bitter W., Welsbeek P.;

```

```

RT "Role for the outer membrane ferric-siderophore receptor PubB in
RL signal transduction across the bacterial cell envelope.";
RL EMBL J. 13:2805-2813(1994).
DR EMBL; X77918; CA54871.1; -.
SQ SEQUENCE 324 AA; 35846 MW; EFPD2338CF87C55C CRC64;

Q52209 Length: 324 May 30, 2002 09:32 Type: P Check: 5380 ..

1 MNGGATGTSIP GEVAEQAMHW HLELOEPAVS AATLAQMSW QAHPLEHHA
51 WORTQVFAQR LREMSPGQR PLAAHALRPQ QSRRTALKOL SLMAAGAGA
101 WYIKDALVQ DMRADYHSRI GEQRRLTLAD GTQVQLNTDS ALNVAFDQQA
151 RLRLVLRGEM LITRPALADS RPLWVDTEHG RLESTLAQFN VRLHGHTQA
201 TYVQGSVALQ PALHAYPIL LGAGEQASEN QOGLARQAV AAVAPAWSQG
251 MLVAQGQPLA AFIEDLARYR RGHACDPAL AGLRVSGTFP LENTDKIIAA
301 VAETLOLEVO HPTRYWTLK PRMA

11AA_SEQUENCE 1.0
ID 09F0G8 PRELIMINARY; PRT; 364 AA.
AC 09F0G8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ENDOGLUCANASE PRECURSOR.
GN ENDS.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RA Michaud P., Belcich A., Courtois B., Courtois J.;
RT "Cloning sequencing and overexpression of a sinorhizobium meliloti
RT endoglucanase gene.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF233448; AAG44364.1; -.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase; 1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
KW SIGNAL.
FT CHAIN 1 27 POTENTIAL.
FT CHAIN 28 364 ENDOGLUCANASE.
SQ SEQUENCE 364 AA; 39832 MW; C927C770AAE832AF CRC64;

Q9F0G8 Length: 364 May 30, 2002 09:32 Type: P Check: 6799 ..

1 MKSSMTRTAR KGMILPLVRA CGLALTLLAT TSQAVLASGT CLRGINLAGA
51 EFGEDGVFG TAYIYPSDET IRYFADKGFN SVRLPFSMSR LQTSLNADFD
101 AAEEERLKDT VRRLREAGQI VVLDPHNYAR YRGELIGSEA VPYEAFADEW
151 SKLSLAFGNQ DGIAFGIMNE PHTMPTQWL TGANVAIAAI RSTGARNLIL
201 VPGNSWSGAH SMWGEDIYGA NGVVMVGKD PLDHYAFEVH QYLDVFSGT
251 KDNCSRAEDA IAAIENYQW LRDNGKRGYL GQFGVPKDEA CVOALSRMVD
301 TVERGRDVMW GWAYWAGDW WPAEALNIQ PTDAGDRPOL QGLSRALSDF
351 SPQAMTCPSL SGGG

11AA_SEQUENCE 1.0
ID 006954 PRELIMINARY; PRT; 460 AA.
AC 006954;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

```

```

DE RFBG PROTEIN.
GN RFBG.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=017 / BIOTYPE EL TOR / SEROTYPE OGAWA;
RX MEDLINE=92212870; PubMed=1372980;
RA Stroehner U.H., Karageorgos L.E., Morona R., Manning P.A.;
RT "Serotype conversion in Vibrio cholerae O1.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2566-2570(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=017;
RX MEDLINE=95309704; PubMed=7540582;
RA Manning P.A., Stroehner U.H., Karageorgos L.E., Morona R.;
RT "Putative O-antigen transport genes within the rfb region of Vibrio
RT cholerae O1 are homologous to those for capsule transport.";
RL Gene 158:1-7(1995).
DR EMBL; X59554; CAA42138.1; -.
DR InterPro: IPR002047; AKH.
DR PROSITE: PS00256; AKH; UNKNOWN_1.
SQ SEQUENCE 460 AA; 52813 MW; A1AA149D3C014937 CRC64;

Q06954 Length: 460 May 30, 2002 09:32 Type: P Check: 8871 ..

1 MTEKKLVVW VPLPPNSSWR GEGIAQTEN IYRNISPERK IEIVSSKHA
51 EMLVGLEKSN PNISVLTGFG RGSSTKKTIG YVSLNEVEKD SLMDLYIAKL
101 PIIPAIFRKV GMVYSQLEYL LSLYIYSHLQ RGRFSSNMC RYWLPTPIIP
151 YTHLLGGEKF VSEFWDPEVFE YNKEEPLTAE YFVKKLSKHF SNASAIITQS
201 RANKDYLETV MGIESSKINV IYNGSPDYSE FKKQSNLSF SEWWSKSEFS
251 GASKKAAFEA LVNHQLNFSV LWRLLTKNV SNRKIVLIST QNRPYKGFQD
301 LFLVLINECL RRDNYDIFT CNVPTKLKER YPSLYERIE VTRVDNYLHA
351 SLYMSDIVL HPSNVEGGLG AYPQYEAASS GKPSLINTGR HYNEMAEEGF
401 DVDLISSNFV NKEETVDKIE KLINSEYMR QNDAINRLK ISWKESASNY
451 ENVFFGNENA

11AA_SEQUENCE 1.0
ID 000620 PRELIMINARY; PRT; 683 AA.
AC 000620;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE OUTER MEMBRANE PROTEIN NOSA PRECURSOR.
GN NOSA.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=91358317; PubMed=1885521;
RX Lee H.S., Abdelal A.H., Clark M.A., Ingraham J.L.;
RT "Molecular characterization of nosA, a pseudomonas stutzeri gene
RT encoding an outer membrane protein required to make copper-containing
RT N2O reductase.";
RL J. Bacteriol. 173:5406-5413(1991).
RC -1- FUNCTION: REQUIRED TO MAKE COPPER-CONTAINING N-2-O REDUCTASE.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE.
CC EMBL; M60717; AAA25906.1; -.
DR HSSP; P05825; 1IEP.
DR InterPro: IPR001005; Myb_DNA_bind.
DR InterPro: IPR000531; TonB_boxC.

```

DR Pfam: PF00593; TonB_box; 1.
KW PROSITE; PS00037; MTB_1; UNKNOWN_1.
FT SIGNAL 1 44
FT CHAIN 45 683
SQ SEQUENCE 683 AA; 74924 MW; 75E652273ADC4CF1 CRC64;

Q00620 Length: 683 May 30, 2002 09:32 Type: P Check: 5873 ..

1 MSSITRGCGA RARISERFSI QPSRLRWQIA HATLLGLLAS GALAAESVDH
51 SEHAHHASSA ELAPMWITGV AQQSPPLTVAT DPKIPRQVP ASDAGDYLO
101 IPGFSAVRGG GSNSDPVFERG MEGSRKLKLLA NGAEMLGACP SRMDSPSSYI
151 TPENYDALTV IKGPQTVLMG PGNSAATILL ERDPEDFSEL GGRIDASFLV
201 GSDGRFDRNI DAAAGGEGGY IRLLANRSDS DDYODGNGDD VHSRWKMWST
251 DLVLGWTPE DFLLELTVGR GDGEARYAGR MMDGSQFERE SVALRFEKTN
301 LGENLKKIEA RYYNYADHV MDNYSLTRPP MMQMATNVDR RTLGGMAAT
351 WOLDEVELVT GYDAQTNEHR RRGVDYKSK PWEKADAFHN YGLFEGELTRT
401 LNDSDRVIGG ARLDHATAKD YRSTGPSAGD SRSDNLPSCF LRYEHLQSL
451 PATAYVGLGH TQRPDPYMWEL FSGGADAFEK LDPEKTQLD FGLQYSKGPL
501 DAWYSAYVGQ VRDYILFSYS PSRYSENIDA RIMGELGAT YRLTSMWKT
551 ASLAYAWGKN SSDGEALPQM PPLEGRIGLT YEOGDWSAAG LMRVVAQNR
601 VAEGKGNVTS KDEDESSGFG VESINGAYRV NQNFKLSTGI DNLEFDKAYSE
651 HLNOAGNAGI GLSADERINE PGRTWMARVD MSP

IIA_SEQUENCE 1.0

ID Q93QE9 PRELIMINARY; PRT; 567 AA.

AC Q93QE9;

DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE PHB SYNTHASE.

GN PHB.

OS Azotobacter vinelandii.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Azotobacter.

OX NCBI_TaxID=354;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-UW 136;

RX MEDLINE-20578894; PubMed-11137297;

RA Segura D., Vargas E., Espin G.;

RT "Beta-ketothiolase genes in Azotobacter vinelandii.";

RL Gene 260:113-120(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-UW 136;

RA Segura D., Espin G.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

DR EMBL; AF267243; AAK72597.1; -

201 NELFOLIOYR PLSEKYOQRP LLIYPPSINK YYILDRPEN SLVRYALEOG
251 HOVELVSWRN EDASGAKTW DNEIQDDAIK AIKYTRAISG GQPLNCVFC
301 IGGTLSTAL AVLEAGSKD HISSLTLLAT FLDYSDTQVI NVEVDEQFVT
351 QRERTIGKG GPVGLFRGQD MGNTFSLRP NELMWNVTVD KYLKGQKPR
401 LDMLFNNDNS TNLPGMYCW YLRHTYLQND LKSGDLECCG VRILDSKIK
451 PLYLLGTQDD HIYFWSAYN TSNLSGKIR FVLGASGHIA GVINPPAQN
501 RHYWTNEQTP ADPDIWETA EKKPGSWND WFAMLVQHAG EOGPAVKQSG
551 NREYQVIEAA PGRYVKG

IIA_SEQUENCE 1.0

ID Q93N64 PRELIMINARY; PRT; 314 AA.

AC Q93N64;

DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE HYPOTHEICAL 36.1 KDA PROTEIN.

OS Coxiella burnetii.

OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;

OC Coxiella group; Coxiella.

OX NCBI_TaxID=777;

RN [1]

RP SEQUENCE FROM N.A.

RA Hoover T.A., Vodkin M.H., Williams J.C., Culp D.W., Thompson H.A.;

RT "A chromosomal DNA deletion explains the phenotype of the Coxiella

burnetii phase II variant.";

RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF387640; AAK71258.1; -

KW Hypothetical protein.

SQ SEQUENCE 314 AA; 36137 MW; EAE20BCDA772D2A7 CRC64;

Q93N64 Length: 314 May 30, 2002 09:32 Type: P Check: 559 ..

1 MKTLVVGIGV QGKKRFFAG NDYGVSDPI NHEADYAGIH EYPLSDYEAA
51 LVCLPDEPKY EVIRYLLSHD KHVLEKPLM TPERKQFEL QKLAQEHNV
101 LYTAYNHRFE PHIKSISELL KENVIGEYR CRLFYNGTA KLVYDSGWRD
151 AGGYLTDLG CHLDLFDYDWM FGLRDSNYVC IDAOCHENNA PDHVVILNRV
201 AKIKIEFEWS LLSWRNSFYC DIIKESGLH IDSICKWGPS QLIIFRQALP
251 SGRPIKTTT LIQSDPTWEL EYRYFKSLIE AKQATNLEKD IMIYDQNLGL
301 TQALKEIDL CTYQ

IIA_SEQUENCE 1.0

ID Q9P981 PRELIMINARY; PRT; 385 AA.

AC Q9P981;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE ENDOGLUCANASE 1.

GN EG 1.

OS Robillarda sp. (strain Y-20).

OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Robillarda.

OX NCBI_TaxID=72589;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Y-20;

RA kashiwagi Y.;

RT "Endoglucanase gene from cellulytic fungi, Robillarda sp. Y-20.";

RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB030819; BAA90480.1; -

DR EMBL; AB030819; BAA90480.1; -

DR EMBL; AB030819; BAA90480.1; -

DR EMBL; AB030819; BAA90480.1; -

DR EMBL; AB030819; BAA90480.1; -

DR EMBL; AB030819; BAA90480.1; -

DR pfam; PF00150; cellulase; 1.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN.1.
SQ SEQUENCE 385 AA; 41428 MW; 149604D42369AD33 CRC64;
Q9P981 length: 385 May 30, 2002 09:32 Type: P Check: 4445 ..
1 MKLVFSALAS LLSGASATIV YAGVAESSGE FGVWSATQTP GTGLPGRGV
51 DYAFISEAAV DVHVDQNHNL LFRVAFLLER MCPPATGLGA AFNETHPDYF
101 KEAVDYITVT KGAYAILDPH NYMRYNDPSY QPFGSGVIGN TSDSTAATTE
151 QFGEFWGELA SRFNDNERVI FGLMNEPHDM ATSLVLANQ AAIDAIRAAN
201 ASNLIIMPEN SWTGGHSWTE GSDPSSALLN QFKDPLNNTA IDIHEYLDYD
251 FSGGHLECVS DPETNLAALT AMLKENNLKA FITEFGGSNS TSCQEMLPDL
301 INYMADNAEY IGTWAMAACP FWGPNSPCCT NSTQLGLEP GSTAVDGSPG
351 LYDTWLPVI QPLVPTELQW SGPASISGGE LTSRA
11AA_SEQUENCE 1.0
ID Q9UW21 PRELIMINARY; PRT; 429 AA.
AC Q9UW21;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HYPOTHETICAL 49.7 KDA PROTEIN.
GN OBPALEHA.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SC5314;
RX MEDLINE=99387068; PubMed=10455055;
RA Hull C.M., Johnson A.D.;
RT "Identification of a mating type-like locus in the asexual pathogenic
RT yeast Candida albicans.";
RL Science 285:1271-1275(1999).
DR EMBL; AF167163; AAD51409.1; -
DR InterPro; IPR000648; Oysterol_BP.
DR Pfam; PF01237; Oysterol_BP; 1.
KW Hypothetical protein.
SQ SEQUENCE 429 AA; 49741 MW; 259C73A904D5795D CRC64;
Q9UW21 length: 429 May 30, 2002 09:32 Type: P Check: 7245 ..
1 MGLTAKLKL KLKTTSEDI DDSGHAKPDN SDDIDEVDSE YQNLISIIA
51 QLRPGMDLSK ITPFTILEK KSMLEIRITNF FQIPKLIDS NSIEDPLDRF
101 IGVLRWYLAS WHISPKAVKK PLNPVLGEVF TCYWDELNN KSAYYLSQI
151 SHHPKSSSYF YIMPEEKIRV DGVVIPKSRF LGNSSAIME GCGYVTLGW
201 DNEVYVMQNP NVYVIGILEG KMRTLEGDM YKCCERNGLE ANIEFKTGF
251 IYGTYDAIEG IIKDSETOKE LFQISGKNE VMYIKNIKTG KKEVLYDTRG
301 SKTLKPKVRP LEEQWDFESR KLMKPTIAGL AKRNHELATE ESKVENEQR
351 IKAKKRLSDG VEFHPKFFRE VNENDGVKN LEVYIKKFD LKEDPEVLE
401 RLFTVAPIVP GQKFEKFFHY PAFKKPESN
11AA_SEQUENCE 1.0
ID Q9UI23 PRELIMINARY; PRT; 190 AA.
AC Q9UI23;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PRO0529.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Zhang C., Yu Y., Zhang S., Ouyang S., Luo L., Wei H., Zhou G.,
RA Zhang Y., Liu M., He F.;
RT "Functional prediction of the coding sequences of 9 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF11848; AAF16687.1; -
SQ SEQUENCE 190 AA; 21480 MW; 4B8104A29AA33844 CRC64;
Q9UI23 length: 190 May 30, 2002 09:32 Type: P Check: 6679 ..
1 MCLCVCLYVC ICVYVCVCHF VCFWVCLPLR VSVYLYLRVC VCVCVFVCLC
51 MCVRCVSVVC VCVCIEREGE RKGATDGSAN KYVPHSQPWE ESNPPTGQD
101 QLWMCCLADSG NVTFLRMGL HFLGKECRSW SLKECFEFPF VIERAQPCVH
151 WLTVTNLRVG DSHRETEGT ADSEQESGCT SLPLGNPOL
11AA_SEQUENCE 1.0
ID Q9NZH3 PRELIMINARY; PRT; 280 AA.
AC Q9NZH3;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE C-TYPE LECTIN-LIKE RECEPTOR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20135876; PubMed=10671229;
RA Colonna M., Samaridis J., Angman L.;
RT "Molecular characterization of two novel C-type lectin-like receptors,
RT one of which is selectively expressed in human dendritic cells.";
RL Eur. J. Immunol. 30:697-704(2000).
DR EMBL; AF200949; AAF36830.1; -
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
KW Receptor; Lectin.
SQ SEQUENCE 280 AA; 32039 MW; 123C2C32FD5F8F6F CRC64;
Q9NZH3 length: 280 May 30, 2002 09:32 Type: P Check: 2970 ..
1 MQAKYSSTRD MDDDDGDTM SLHSQASATT RHPEPRRTEH RAPSSTWRPV
51 ALTLTLCLV LLIGLAAMGL LFPQYQLSN TGQDTISQME ERLGNTSQEL
101 QSLQVQNIKL AGSLQHVAEK LCRELYNKAG AHRCSPTETQ WKWHGDNQCYO
151 FYKDSKSWED CKYFCLSENS TMLKINKQED LEFAASQSYS EFFYSWTGL
201 LRPDSGKAWL WMDGTPFTSE LPHIIDVYS PRSRDCVAIL NGMIFSKDCK
251 ELKRCVCERR AGWVKPESLH VPPELTGEGD
11AA_SEQUENCE 1.0
ID Q9H245 PRELIMINARY; PRT; 531 AA.
AC Q9H245;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE C10RF28.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2118927; Pubmed=11318611;
RA Sood R., Bonner T.I., Malakowska I., Stephan D.A., Robbins C.M.,
RA Connors T.D., Morgenbesser S.D., Su K., Faruque M.U., Pinkelt H.,
RA Graham C., Baxevanis A.D., Klinger K.W., Landes G.M., Trent J.M.,
RA Carpten J.D.;
RT "Cloning and characterization of 13 novel transcripts and the human
RT RGS8 gene from the 1q25 region encompassing the hereditary prostate
RT cancer (hpc1) locus."
RL Genomics 73:211-222(2001).
DR EMBL; AF312865; AAG45339.1; -
SQ SEQUENCE 531 AA; 60615 MW; EBF7BAAA61A5270B CRC64;

Q9H245 Length: 531 May 30, 2002 09:32 Type: P Check: 5880 ..

1 MADVLSVLRQ YNIQKKEIVY KGDEVIFGEF SWPKNVKTNV VVWGTEKEGQ
51 PREYTYLDSI LFLNNVHLS HPVYVRAAT ENIPVVRPD RKDLGLYNG
101 EASTSASIDR SAPLEIGIQR STQVKRADE VLAFAKKPRI EDEECVRLDK
151 ERLAARLEGH KEGIVQTEQI RSLSEAMSVE KIAAIKAKIM AKKRSTIKTD
201 LDDITALKQ RSFVDAEVDV TRDIVSREKV WRTRTIILQS TGNFSKNIF
251 AILQSVKARE EGRAPEQRP PNAAPVPTL RTKQIPPAV NRYDQERFKG
301 KEETEGFKID TMGTYHGMVL KSVTEGXSAR KTQTPAQPV PRVSGARPP
351 PNOKKGRTP IIPPAATTS LITMLNAKDL LQDLKFVPSD EKKQCGORE
401 NETLIQRRKD QMOPGTPAIS VTVPYRVVDQ PLKLPQDWD RVAVAVVQGP
451 AMQFKGWPWL LPDGSVPDIF AKIKAFHLKY DEVRLDPNVQ KMDVTLELS
501 YHKRHLDRPV FLRFWETLDR YMVKKHSHLR F

11AA_SEQUENCE 1.0
ID Q9BXJ8 PRELIMINARY; PRT; 343 AA.

AC Q9BXJ8;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TRANSMEMBRANE PROTEIN INDUCED BY TUMOR NECROSIS FACTOR ALPHA.
GN TMPIT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Murakami T., Matakai C., Hamakubo T., Kodama T.;
RT "Endothelial cell transmembrane protein induced by tumor necrosis
RT factor alpha (TMPIT) mRNA, complete cds."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327923; AAK16442.1; -
SQ SEQUENCE 343 AA; 40610 MW; 7A6E241804F59A96 CRC64;

Q9BXJ8 Length: 343 May 30, 2002 09:32 Type: P Check: 1303 ..

1 MOPPPGPIG DCLRDWEDLQ QDFQNIQETH RLYRLKLEEL TKLQNNCTSS
51 ITRQKKRIQE LALALKKCRP SLPAEAEGAA QELENNOMKER QGLFEDMEAY
101 LPKKNGLYLS LVLGWVNTL LSKQAKFAVK DEYEKFLYL TIILLISFT
151 CRFLNSRVT DAFNELLWV YCTLTIRBS ILINNGSRIK GWWVFHHYVS

201 TFLSGVMLTW PDGLMYQFR NOFLSFSMQ SFVQFLOY Y QSGCLYRLRA
251 LGERHTMDLT VEGFQSWMR GLTFLLPFL FGHFWQLENA LTLFNLAQDP
301 QCKEWQVLMC GPELLLELG NEFTTLRVH HKFHSQRHGS KKD

11AA_SEQUENCE 1.0
ID Q9NV66 PRELIMINARY; PRT; 732 AA.

AC Q9NV66;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CDNA FLJ10900 F15, CLONE NT2RP5003522, WEAKLY SIMILAR TO
DE NADPH-CYTOCHROME P450 REDUCTASE (EC 1.6.2.4).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RA Isoqal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K.,
RA Masuko Y., Kanehori K.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001762; BAA91891.1; -
DR HSSP; P16435; 1B1C.
DR InterPro; IPR001094; Flavdxn-like.
DR InterPro; IPR001226; Flavodoxin.
DR Pfam; PF00258; flavodoxin.1.
DR PRINTS; PR00369; FLAVODOXIN.
SQ SEQUENCE 732 AA; 83720 MW; F8C92411EC2B239C CRC64;

Q9NV66 Length: 732 May 30, 2002 09:32 Type: P Check: 4753 ..

1 MDPSADTWDL FSPILSLWIN RFIYLGFAV SISLMICVOI VIKTOGRNLQ
51 EKSVPKAAQD LMTNGYVSLQ EKDI FVSGVK IFYGSQGTGA KGFAIYLAEA
101 VTSLDLPVAI INLKEYDPDD HLEEVTSKN VCVFLVATYT DGLPTESAEM
151 FCKWLEEASD DFRFGKTYLK GMRYAVFGLG NSAYASHPNK VGKNVDKMLW
201 MGAHRVMSR GEGDCDVYKS KHGSTEADPR AWKTKFTISOL QALQKERKK
251 SCGGCHCKKGK CESHQHGSEE REEGSHEDDE LHHRTDEEBE PFESSSEEEF
301 GGEDHOSLNS IVDVEDLGI MDHVKKKKRE KEQOEKSGL FRNMGRENEDG
351 ERRAMITPAL REALTKQGYQ LIGSHSGVKL CRWTKSMLRG RGGCYKHIFY
401 GIESHRCMET TPPLACANKC VFCWRHHTNP VGTEWRWMD QPEMILKEAI
451 ENHQMNIKQF KGVPGVKAER FEEGMYVKGH ALSLVGEPIM YPEINREPLK
501 LHQCKISSFL VTNAQPAEI RNLEPTQLY VSVDASTRDS LKKIDRPLFK
551 DFWQRFDSL KALAVKQQRV YRRLTLVKAW NVDELQAYAQ LVSLGNPDFI
601 EVKGVTYGGE SSASSLTMAH VPWHEEVQF VRELVDLIFE YEIACEHSHS
651 NCLLIAHRKF KIGGEWWTWI NYNRFOELIQ EYEDSGSKT FSAKDYMART
701 PHWALFGASE RGFDPKDTRH QRKNKSAIS GC

11AA_SEQUENCE 1.0
ID Q9H9C4 PRELIMINARY; PRT; 561 AA.

AC Q9H9C4;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE CDNA FLJ12855 FIS, CLONE NT2RP2003506, WEAKLY SIMILAR TO
DE NADPH-CYTOCHROME P450 REDUCTASE (EC 1.6.2.4).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fuji A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuko Y., Oshima A.;
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK022917; BABI4307.1; .
DR HSSP; P16435; 1B1C.
DR InterPro; IPR001226; Flavodoxin.
DR Pfam; PF00258; flavodoxin; 1.
SQ SEQUENCE 561 AA; 64682 MW; 4DF7641582E575F8 CRC64;

Q9H9C4 Length: 561 May 30, 2002 09:32 Type: P Check: 6460 ..

1 MRYAVFGLGN SAYASHENKV GKNVDKWLMM LGAHRVMSRG EGDGCDVVKSK
51 HGSIEADFERA WKTKFISQLQ ALQKGERKKS CGCHCKKKGK ESHQHGSEER
101 EEGSHQDEL HHRDTEEEP FESSSEEEFG GEDHQSLNSI VDVEDLGKIM
151 DHVKREKREK EQQEEKSGLF RNMGRNEDGE RRAMITPALR EALTQGYQL
201 IGSMSGVKLC MWTSMERGR GGCYKHTFYG IESHRCMETT PSLAYANKCV
251 FCWRHHTNPV GTEWRMKMDQ PEMILKEAIE NHQNMKQFK GVPGVKAERF
301 EEGWYVKHCA LSLVGEPIMY PEINRFKLKL HQCKISSFLV TNAQFPAPFIR
351 NLEPYTQLYV SVDASTKDSL KRIDRPLEKD FWOQFLDSLK ALAVKQORTV
401 YRLTLVKAWN VDELQAYYQL VSLGNPDFIE VKGVTYCGES SASSLTMAHV
451 PMHEEYVQFV RELVDLPEY EIACEHEHSN CLIAHRKEK IGGEWWTWIN
501 YNRFQELIQE YEDSGGSKTF SAKDYMARTP HMALEGASER GFDPKDTRHQ
551 RKNKSKAISG C
11AA SEQUENCE 1.0
ID Q9H5L7 PRELIMINARY; PRT; 355 AA.
AC Q9H5L7;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
DE CDNA: FLJ23316 FIS, CLONE HEP12031.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026969; BABI5608.1; .
SQ SEQUENCE 355 AA; 40699 MW; 84ABF38A6EB2E7C8 CRC64;

Q9H5L7 Length: 355 May 30, 2002 09:32 Type: P Check: 1905 ..

1 MSVEKIAACNQ TKIMAKKRST IKTDLDDIT ALKORSFVDA EVDVTRDIYS

51 RERWRTTIT ILQSTGKNES KNIFAILQSV KAREGRAPE QRPAPNAPV
101 DPTLRTPQPI PAAYNRYDQE RFKKEETEG FKIDTGTGYH GMTLKSVTGEG
151 ASARKTQTPA AQPVPRPVSO ARPPNQKKG SRPTIIIPA APTSLKTMLN
201 AKDLLODLKF VPSDEKKKG CORENETLIO RRDQMOPGG TAISVTPYR
251 VVDQPLKMP QDWDRYAVF VQGPAMQFKG WPMLLPDGSP VDIFAKIKAF
301 HLKXDEVRLD PNVQKWDVTY LELSYHKRHL DRPVFLRFE TLDRTYVKKHK
351 SHLRF
11AA SEQUENCE 1.0
ID Q96KN6 PRELIMINARY; PRT; 1449 AA.
AC Q96KN6;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE PUTATIVE TRP CATION CHANNEL.
GN KNP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wehage E., Eisfeld J., Heiner I., Juengling E., Zitt C., Lueckhoff A.;
RT "Splice variants of LTRPC2 differentially activated by ADP-ribose an
RT hydrogen peroxide."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ417076; CAD01139.1; .
SQ SEQUENCE 1449 AA; 165328 MW; 563DAB259F55A9E7 CRC64;

Q96KN6 Length: 1449 May 30, 2002 09:32 Type: P Check: 6589 ..

1 MEPSALRKAG SEQEGFEGF PRRVTDLGMV SNLRSSNSL FKSRLQCPF
51 GNNDKQESLS SWIPENIKKK ECVYFEVSSK LSDAGKVVCQ CGYTHEQHLE
101 EATKPTFQG TQWDEKHHVQ EMPTDAFGDI VETGLSQVKV KYRVVSQDTP
151 SSVIYHMTQ HMGLDVNNLL ISVTGAKNF NMPRLKSTF RRGVLKVAQT
201 TGAWITGGS HTGVAKQVGE AVRPDSLSSS YKRGELITIG VATWGTVHRR
251 EGLHPTGSF PAEYILDEDG QGNLTCLDN HSHFILVDDG THGQYGEIIP
301 LRTLRLEKFI EOTKRGGVA IKPIVCVVL EGGPGTLHTI DNATNGTPC
351 VVVEGSGRVA DVIAOVANLP VSDITISLIQ QKLSVEFQEM PETTESRIIV
401 EWTKTIQDIV RRRQLLTVFR EKGDGQQDQD VALIQALLKA SRSQDHFGHE
451 NMDHQKLAV AMNRVDIARS EIRNDEWQWK PSDLHPTMTA ALISNKPEFV
501 KLELENGVQL KEFVTWDTLL YLYENLDPSG LFHSKLQMHM VAQVLRRLG
551 DFTQPLPRP RHNDRLRLLL PVPHVKNVQ GVSLSLYKR SSGHYTFTMD
601 PIRDLIIWAI VQNRRELAGI IWAQSQDCIA AALACSKILK ELSKEEDTID
651 SSEMLALAE EYEHRAIGVF TECYRKDEER AQKLLTRVSE AWGKTTCLQL
701 ALEAKDMKFV SHGGIOAFLT KVMWGQLSVD NGIMRVTLICM IAPPLLTGL
751 ISFREKRLQD VGTPARARA FETAPVVVFH INILSYFAFL CLFAYVLMDV
801 FQPVSWCEC AIYVLMFSLV CEEMRQLFYD PDEGLMKKA ALYFSDFWNK
851 LDVGAILLFV AGLCRLIPA TLVPGRVILS LDFILCLRL MHIETISKTL

901 GPKIIIVKRM MKDVFFELFL LAWVVSFGV AKQAILLINE RRVDMLEFGA
951 VVHSYLTIFG QIPGYIDGVN FNDHCSPNG TDPYKPKCE SDATQORPAF
1001 PEWLVLLLC LVLLFTNILL LNLIAMFNY TPQOVQEHND QIWKFORHDL
1051 IEEYHGRPA PPPFILLNHL QLEIKRVVLK TPAKRHKQLK NLEKNEEA
1101 LLSWEIYLKE NYLQNRQFOQ KORPEQKIED ISNKKVDAMVD LLDLDPLKRS
1151 GSMEORLASL EEOVAQTARA LHWIVRTLRA SGRSSADVP TLASQAAEE
1201 PDAERGGRKK TEEPDSYHV NARHLLYPNC PYTRFPVPNE KYPWETEFLI
1251 YDPEYTAER KDAAMPDMG ENPMGRTGLR GRGSLSCFGP NHTLYPMVTR
1301 WRNEDGAIC RKSIIKMLEV LVVKLPLSEH WALPGSREP GEMLPKLR
1351 ILRQEHWPSE ENLKCMEV YKGYMDPRN TDNAWETVA VSVHFQDQND
1401 VELNRLNSNL HACDSGASIR WQVVDRIPL YANHKTLLQK AAAEFGAHY

!!AA_SEQUENCE 1.0
ID Q96JM4 PRELIMINARY; PRT; 1227 AA.
AC Q96JM4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE KIAA1801 PROTEIN (FRAGMENT).
GN KIAA1801.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE-21245130; PubMed-11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
DR EMBL; AB058704; BAB47430.1; -.
FT NON_TER 1 1
SQ SEQUENCE 1227 AA; 142754 MW; A20651382DC61B8F CRC64;

Q96JM4 Length: 1227 May 30, 2002 09:32 Type: P Check: 5997 ..

1 RTREKQOEK EEFMRSTDC ATPDFVPEPS PHDLPMDHV LPDDADINFG
51 YCEVEEKCQ SFEAWQEKQ ELDEKKEQTL KAQRDREKQ FQEEEEKRHC
101 WMKQKVEKK KLENIOKQEQ DKMNDLYKE EKIMKEKFKQ HEEYIRNLHL
151 QMEERTRFK DQEKKEKNSL LKQONNAVK IQAKYKAFVA YQKYGPITKE
201 QIESKKRKAQ EWKEEKAIR QKEENRRL EEEQRIKEER KKQKEEERKR
251 REKEYEKKK IYQEREQLI SKEKILRED ASQQLISSA LKSGYNKH
301 LSLDISNDK GDIAKNLVE NSKKQEDVLL WLYEESNMKE NYDRQTLIKE
351 SIQVLRKESI SSQTLADFK MEKENENLAK KRCSEELVKQ ERKYENTDNK
401 TELGNSDLKG NLKEQFPLQE LKSDAQKEEK IMKHVINENT GOKTQIILGH
451 NQETSEVKTN EEQKIKNQ OKKIQKVEKE EIQEQNGLLY KDKDTLVISV
501 KQSLSLTSE NSKDVRENVI LQEKELYSKS KETENPKDN AMNSGIIVFN
551 TTDMINIEG KRNDQDYLQ RHAPCEGLSN YNASSMVSX EVNSIKSEIR

601 NISEKCHENA PEPSMTCCV SESTLLYSIE ERLAWIKSF KPWLEIFKQ
651 QOKKIYRRKR PVKCRANMP ALDKLEILRC GPWDTLQOVT TYTFQDLPGC
701 VLSTIAECTN LQFLSLRRCG LTSLSLSLNC KKLKYIDAQ NHIEAIECEN
751 LENLCVLLN KNQLSLHGL DGCINIQCLE LSYNKITRIG YSFLEEKLV
801 DNTGFGHLG TSTSYLSLAQ WPIPTGLCWS WIPITSLTKN SDCNPLISHL
851 YWNCGLSELK NLOQLLDHN QLINTKGLCD TPPIVYLDOS HNLDTVEGV
901 ENCGLLQILK LQGNILSELP SLENVLLRE LHLDDNSIST VEAFFSSYMLP
951 LLQNTISQN SLTKIVPLFH FVSLKLDVS HNCUSDLSKA IKWFDACYSL
1001 HELSLGNPL LQETWRDSL LKVLPAIRIL NGNLINSNSE SRTEEHNLQ
1051 SAGFLALCQ QIREPNLLIE NYITGKGVF TLDTAENLCH YFKKIMILST
1101 EYRHAHERGD VTIITKDESE AOKNHLAPT N SDSTLQNGVF YSCAREGEPP
1151 SPDIEKAMD SVSSHPLSK SATCENMEGR HOELVLCQKR EDSKASSIPT
1201 IRIPREYVM TNSLLRNHON IEPSEKM

!!AA_SEQUENCE 1.0
ID Q96A07 PRELIMINARY; PRT; 355 AA.
AC Q96A07;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SIMILAR TO HYPOTHETICAL PROTEIN FLJ23316.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BLADDER CARCINOMA;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-LUNG CARCINOMA;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014351; AAH14351.1; -.
DR EMBL; BC011808; AAH11808.1; -.
KW Hypothetical protein.
SQ SEQUENCE 355 AA; 40621 MW; 3F228230AB6F3E9F CRC64;

Q96A07 Length: 355 May 30, 2002 09:32 Type: P Check: 1525 ..

1 MSVERIAAIK AKIMAKKRST IKTDLDDIT ALKQSFVDA EYDVTRDIYS
51 RERVWRRTT ILQSTGKNFS KNIPAILQSV KAREEGRAPE QRPAPNAPV
101 DPTLRKQPI PAAVNRVDQF RFKGRKETEG FKIDTMGTYN GMTLKSVTG
151 ASARKTQTPA AQPVPRVSQ ARPPNQKKG SRTPILIIIPA ATSLITMLN
201 AKDLLQDLKF VPSDEKKKG CQRENEFLIQ RRDQMOPGG TAISVTVPYR
251 VVDQPLKMP QDWDRVAVF VQGPAAQFKG WPMLLPDGSP VDIFAKAKAF
301 HLKYDEVRLD PNVQKWDVTV LELSYHKRHL DRYVLRFEW TLDRYMKKH
351 SHLRF

!!AA_SEQUENCE 1.0
ID Q9V696 PRELIMINARY; PRT; 2040 AA.
AC Q9V696;

DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE CG8487 PROTEIN.
GN CG8487.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler C., Cadieu E., Center A., Chandra I.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003822; AAF58532.1; -.
DR HSSP; Q99418; 1PBV.
DR FlyBase; FBgn0033714; CG8487.
DR InterPro; IPR000904; Sec7.
DR Pfam; PF01369; Sec7; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PS50190; Sec7; 1.
SQ SEQUENCE 2040 AA; 226981 MW; A59107B06813FD7C CRC64;

Q9V696 length: 2040 May 30, 2002 09:32 Type: P Check: 9003 ..

1 MALPNCGIYV VRGEMATIMT AMRRGTRWNA TAYVDDEDS LKLFLIDLKH
51 ELNRIEDLRQ IEPQVFLAPF LEVIRTDAT GPLTSLALAS VNKLLSYGLI
101 DPTSPNLADI VERIADAVTH ARFMGTDOSS DGVTFMRYIE VLHTLIRSPE
151 GAAVSNVSMC EVMLSCEFKIS EEPRLSELLR RSAEKSLEDM VLLFFMRLDQ
201 FAERSDITML QKRFTIGDAA SGATQEKLR KTVAAQATAP RKSSAVEEPP

251 QTPQSANLTV PGHLKAPILA TTPASPAGNI LDMQKITQT PTTASTGED
301 EFTVPEPVI QVESTSEPL LDGETGEATS TLAANSSEY INSVGVRFTQ
351 QSTDHDVTSI SPYGLPFIQE LFRLLILCN PUDKQNSDSM MHTGLSLTFV
401 AFEVAADNIG KYEGELLELVK DDLCRNLISL ISSERLSIFA ADLQCLFLE
451 ESLRGLKFQ LEAVLRKISE IIASDNPKTP YEMRELALDN LLQLMRIPGF
501 VTELYINYDC DLYCTDMFES LTNLLSKYTL SATNAVYSTH IISMDTLISV
551 IDSIERNCAA SKNSSNNRES LPEAPATGG SRHSRHNSGL EGIVDSGNS
601 VAAEKEVENI ASFINASSHR LRLQSGEGV GITSEQLAKV KQKKRLSOG
651 TERNQRPKEK GIOYLQEHGI LNAELDPMV ALFLRENPGI DKKMIGEYIS
701 KKKNVDSKIL INFVDSFDT GLRVDQALRL YLETFRLPGE APLIFVLEH
751 FSDHWHKONQ DPEANVDAF RLAYAIIMN MDQHSNAKR LNVPTLDEDF
801 TKNLRLGLNG EDFDQMLAQ VENAICKNEI VMPAEOTGLV RENYQWKVL
851 RRGDTHDGHF HYVHDASYDV EIFNIVWGS LSALSEFMFK STETGYQRTL
901 AGFSKSAIS AHYNLHSDFD ALVLTICKFT TLLSVEQHE PAPANNETQ
951 AVNFGNGKA QAMRTVELL VHDYGDCLRE SWKHILDLVL QLFRLKLPR
1001 SLIEVEDFCE ANGKAMILE KPREKQESGL FSSLSYFISS EGOREPTYEE
1051 QDFIKLGRKC IKECQLDQML QESKFVQLES LQELLCVLA LLKAPQGHKS
1101 IGLPYAEDQT VFWMEFLVKI VVHNRDRMIP LMPAVRDQMY LLLMGASCG
1151 YDYLLNRCIV AVLKLAIYLM RNEELCPYVL QSLKMLLMK PALLLRISKQ
1201 ISIGIYELK TSAQNIHSEQ DWQIIFNLLE CVGAGAVPPN YDDAQPLPR
1251 NSAKSDGAI SCGEDATAVP ERGYTSDSEI TKASAPAVS SPSAENWILV
1301 NNKDSLTTA SRQSPSLS APPVNTLVYN CQLLDHAPFA LEKWDLSLAF
1351 IVRSVAHITP YNFEACVRCI RIFVEACRDG GIRQRKLES AAKQKSKKR
1401 SERKPGMASS ASSSNLTLLT GDPDNDQING NAAEQEDLAQ RYEQLSIQL
1451 DLMYTLTYRT AQIFRWAAEE GCTVPQSAAL WSPGWCPLIQ GIARLAMDR
1501 REYRTHAISC LOORALLVHD LQTLSTGTEWC SCFHQVLFPL LNELPESNA
1551 AGOLDALLLE ESRIRATIM SKVELQHLLT LIELGNAPNE LMLDILDYIE
1601 REMKVGSDTL SEQMQEILKN MLLVMHSVRV FHNQDGSLOQ ALMELTWRI
1651 GEFLPNLKEE LFHDEDSITP AVSLDYTQIR FAHSTGTTDY SLSLPLYGYS
1701 PKISSISASK CDTLPATQK KPWAKLRSL KLLKKAAYAA APOQQLPAVT
1751 ILPRQTOVSN ELVVSAPTTP AATPLLGSPV ESPRSIILQ PMADVLAQP
1801 PSFVEAQPII VPPQPAVTD PIPSTLLPD LVNEATAAV QATTSPTHS
1851 PQEAQOPASI VOQTNIVTNN NTYNSYAIEV PMAPETTAEQ FGQOQOQLLY
1901 QOYYOQYQAO QOOLPAPASD PAIINPISHL LAGNAYPSLP KMPQASIVHS
1951 FAPVYESQAA TSGAGTAABD TYOEYVQNPY NLTLOOHPOQ QLHQOQOQOQ
2001 QOATGMANAF PAVATPANYF NVNVDPSISIP PGSELLYGQO

11AA_SEQUENCE 1.0

```
ID 09V9K3      PRELIMINARY;      PRT;      337 AA.
AC 09V9K3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CG14593 PROTEIN.
GN CG14593.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotilier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003784; AAF57285.2; -.
DR FlyBase; FBgn0033058; CG14593.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS50262; G_PROTEIN_RECIP_F1_2; 1.
SQ SEQUENCE 337 AA; 36989 MW; E06E7260DC3249C4 CRC64;
```

09V9K3 Length: 337 May 30, 2002 09:32 Type: P Check: 7480 ..

```
1 MCRISSEFKD ISIGSVFTL TALSGERYCA IVNPLRLQT KPLTVFTAVM
51 IWLAILLGM PSVLSFDIKS YPVFTATGM TIEVCSFPRD PEYAKEMVAG
101 KALVYLLPL SIIGALYIMM AKRLHMSARN MPEGOQSMOS RTQARARLHV
151 ARWVAFAVV PFICFFPYHV FELWYHFYPT AEEDDEFWN VLRIVGCTCS
201 FLNSCVNPVA LYCVSGVERQ HENRYLCIC VKRQPHLRQH STATGMNDT
```

```
11AA_SEQUENCE 1.0
ID 09GN87      PRELIMINARY;      PRT;      209 AA.
AC 09GN87;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE GLUTATHIONE TRANSFERASE GST1-1 (EC 2.5.1.18).
GN GST1.
OS Anopheles dirus, and
OS Anopheles dirus B.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Anopheles.
OX NCBI_TaxID=7168, 123217;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=A.dirus;
RA Ketterman A.J., Prommeenate P., Boonchaey C., Chanama U.,
RA Leetachewa S., Promtet N., Prapanthadara L., a.;
RT "Single Amino Acid Changes Outside The Active Site Significantly
RT Affect Activity of Glutathione S-Transferases.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=A.dirus B;
RX PubMed=11102837;
RA Pongjaroenkit S., Jirajaroenrat K., Boonchaey C., Chanama U.,
RA Leetachewa S., Prapanthadara L., Ketterman A.J.;
RT "Genomic organization and putative promoters of highly conserved
RT glutathione S-transferases originating by alternative splicing in
RT Anopheles dirus.";
RL Insect Biochem. Mol. Biol. 31:75-85(2001).
DR EMBL; AF273041; AAG38507.1; -.
DR HSSP; P30712; 1LJR.
DR InterPro; IPR004046; GST_C.
DR Transferase.
SQ SEQUENCE 209 AA; 23397 MW; A555F1268441DAD6 CRC64;
```

09GN87 Length: 209 May 30, 2002 09:32 Type: P Check: 9284 ..

```
1 MDFVYLPGSA PCRAVQMTAA AVGVELNLKL TNLMAGEHMK PEFELKLPQH
51 CIPFLDNGF SLWESRAIQI YLVEKYGKDD KLYPKDPQKR AVYVQRLFFD
101 MGTLYQREGD YWPQIFAKQ PANANEKKM KEAVGELNLF LEGQEYAAGS
151 DLTIDLUSLA ASIATYEVAG FDFAPYPNVA AWLARCKANA PGYALNQA
201 DEFRAKEMS
11AA_SEQUENCE 1.0
ID 09NL42      PRELIMINARY;      PRT;      491 AA.
AC 09NL42;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE BRACHYURY.
GN CSBRACHYURY.
OS Clona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Clonidae; Clona.
OX NCBI_TaxID=51511;
RN [1]
RP SEQUENCE FROM N.A.
RA Imai K., Takada N., Satoh N., Satou Y.;
RT "An essential role of beta-catenin in the endoderm specification of
```

```
RT ascidian embryo."
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB031545; BAA92187.1; -.
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
SQ SEQUENCE 491 AA; 54673 MW; 62E83F723AA03025 CRC64;

Q9NL42 length: 491 May 30, 2002 09:32 Type: P Check: 320 ..

1 MTSTLESCLM TGQNTNESSC VKMSLVDQNL WSRFHAFVNE MIVTKNGRRM
51 FPLVLTITISG LDPTAMYSVM LDFVAVDNNR WKYVNGEWP GKKPEPHYTS
101 CAYIHPDSPN FGSWMMKQPV GFSRVKLTNK ATGNAQQIML NSLHKYEPR
151 HIMKVGGAES QQIVATHSFA ETRFIAVTAY QNEDVTSIKI KYNPEAKAFL
201 DAKESRTETE TYKKTANAG TSQTYARVSS WTPGVNQTNG DNLPSVYHCD
251 PATRNDFGVN PMFAPTNCPS PYQYARERR SSRSQSRSHN YHPYSHRDAH
301 HPMPIHQDIS TSIDVNATPS HETLCYPTGY QVMTADQWPH TAIPHHDVAE
351 QGTESLDQVS VEDMASLAFE SATSSASLTQ SDFSTVDSFD SSSFGAHMSY
401 STNHIGYNS NSPISEFERN AHTAHYIPNP YEMSPQDCLN VAATEENITIT
451 ASLYHSPSP SAELGQSSPL ADAYDPTKLT SNWTPLTPPS L

11AA_SEQUENCE 1.0
ID Q9BLF5 PRELIMINARY; PRT; 308 AA.
AC Q9BLF5;
DT 01-JUN-2001 (Tremblrel. 17; Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE BRACHYURY PROTEIN (T) (FRAGMENT).
GN OILBRA-U.
OS Oikopleura longicauda.
OC Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
OC Oikopleuridae; Oikopleura.
OX NCBI_TaxID=107032;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21347418; PubMed=11455437;
RA Nishino A., Satou Y., Morisawa M., Satoh N.;
RT "Brachyury (T) gene expression and notochord development in Oikopleura
longicauda (Appendicularia, Urochordata).";
RL Dev. Genes Evol. 211:219-231(2001).
DR EMBL; AB052730; BAB33364.1; -.
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
FT NON_TER 308
SQ SEQUENCE 308 AA; 35259 MW; 7FD78FE87C6FBA86 CRC64;

Q9BLF5 length: 308 May 30, 2002 09:32 Type: P Check: 1820 ..

1 MNGTSDIKIE DHHNHEKDN I KVRLEDEELM KKFSSLTNEM IYTKNGRRMF
51 PYMKVRVGGI DENAMYSVL DFSAADNHRW KYVNGEWP GKKPEQAPPC
101 VYMHDPSPNF GSHWMKQSIG FSKVKLTNKL NNHNGSQIML HSLHKYEPR
101 VYMHDPSPNF GSHWMKQSIG FSKVKLTNKL NNHNGSQIML HSLHKYEPR
```

```
151 HIIVGGQGG TQOILKTQTF PTKFIAVTA YONEITSIK IKHNPEAKAF
201 LDKAKOREID QNNQNNWYOV PIENNPTPTQ SRKASRNKOS CRSQPYEIKR
251 NONHSSNHL PSLPQSNQIE PTYPAYEPY SPQMDPMAPS GAGTPPLASI
301 GQMLDWNS

11AA_SEQUENCE 1.0
ID Q9BLF4 PRELIMINARY; PRT; 403 AA.
AC Q9BLF4;
DT 01-JUN-2001 (Tremblrel. 17; Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE BRACHYURY PROTEIN.
GN OILBRA.
OS Oikopleura longicauda.
OC Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
OC Oikopleuridae; Oikopleura.
OX NCBI_TaxID=107032;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21347418; PubMed=11455437;
RA Nishino A., Satou Y., Morisawa M., Satoh N.;
RT "Brachyury (T) gene expression and notochord development in Oikopleura
longicauda (Appendicularia, Urochordata).";
RL Dev. Genes Evol. 211:219-231(2001).
DR EMBL; AB052731; BAB33365.1; -.
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
SQ SEQUENCE 403 AA; 45610 MW; 86A545378669B6EE CRC64;

Q9BLF4 length: 403 May 30, 2002 09:32 Type: P Check: 7868 ..

1 MNGTQDIKIE DHHNHEKDN I KVRLEDEELM KKFSSLTNEM IYTKNGRRMF
51 PYMKVRVGGI DENAMYSVL DFSAADNHRW KYVNGEWP GKKPEQAPPC
101 VYMHDPSPNF GSHWMKQSIG FSKVKLTNKL NNHNGSQIML HSLHKYEPR
151 HIIVGGQGG TQOILKTQTF PTKFIAVTA YONEITSIK IKHNPEAKAF
201 LDKAKOREID QNNQNNWYOV PIENNPTPTQ SRKASRNKOS CRSQPYEIKR
251 NONHSSNHL PSLPQSNQIE PTYPAYEPY SPQMDPMAPS GAGTPPLASI
301 GQMLDWNSG GYYESHDNNS VOLNGSPTSS HLSFNDGIQG TMTNATSGAN
351 GASSTLPPIN DFATQYNSY NDYALNDHYS IDPTQSPDWS TSPNLALAWA
401 ANL

11AA_SEQUENCE 1.0
ID Q17041 PRELIMINARY; PRT; 491 AA.
AC Q17041;
DT 01-JAN-1998 (Tremblrel. 05; Created)
DT 01-JAN-1998 (Tremblrel. 05, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE HYPOTHEICAL 56.5 KDA PROTEIN.
GN T15B7.16.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode *C. elegans*: a platform for
investigating biology. The *C. elegans* Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Pauley A., Gattung S.;
RT "The sequence of *C. elegans* cosmid T15B7.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL; AF022985; AAB69969.1; -;
DR InterPro; IPR000188; GABA_receptor.
DR InterPro; IPR001175; Neur_channel.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_membr; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Glycoprotein; Hypothetical protein; Ionic channel;
KW Postsynaptic membrane; Transmembrane.
SQ SEQUENCE 491 AA; 56494 MW; FAFB9922F24F6AEA CRC64;
OI7041 Length: 491 May 30, 2002 09:32 Type: P Check: 8670 ..
1 MEVEKONLKK KLQPNSPRE HSPGEPAND TISQLVLRRL FDDNYKNKTV
51 PLKTSATVVS VEEGIONIAQ VSEISASFTL DLLFSQIWHM PRLRFDLHTN
101 CLQNLTLGYS MVEKLWTPNV CFVNSKKTET HSSPTNIFL MIYPNGTVMV
151 NYRLQVOSPC MVDLVLFPMI IMNCLELIES YAYNAKVKL NWREMPVFS
201 IAKSKLSDFT LYGLQWTKNS FEYAAGQWDQ VSSECSLNCF SKMDGFQLTV
251 SLTFSRAYGF YILQMYIPTV SSVLSFVSF WIDLKALPAR ITLGVSSLMA
301 LTFQYGNVAK NLPRVGVKS IDVYMLTGA FIFLTMIEVA FVCYLDSENN
351 LRRKERQAEK KKRVAVLQR KDKRKKNNYG ATTVTNANAA ESDLRSNYDE
401 PFSQNGTTSK KISISRQANN MFESLHALAQ FGLLTDDDE NTKWTAQNVD
451 KFCRAKAPLS FCFNLNIYWC YLLYQNYLAK AEALAQMTTP S
11AA_SEQUENCE 1.0
ID Q9XYS1 PRELIMINARY; PRT; 438 AA.
AC Q9XYS1;
DT 01-NOV-1999 (Tremblrel. 12, Created)
RT 01-NOV-1999 (Tremblrel. 12, last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE BRACHYURY PROTEIN.
GN BRA.
OS Clona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Clonidae; Clona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97195701; PubMed-9043074;
RA Corbo J.C., Levine M., Zeller R.W.;
RT "Characterization of a notochord-specific enhancer from the Brachyury
promoter region of the ascidian, *Clona intestinalis*.";
RL Development 124:589-602(1997).
RN [2]

RP SEQUENCE FROM N.A.
RA Corbo J.C., Levine M., Zeller R.W.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF123247; MAD21079.1; -;
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
SQ SEQUENCE 438 AA; 49625 MW; 1DDFCB65C2D92F4F CRC64;
Q9XYS1 Length: 438 May 30, 2002 09:32 Type: P Check: 646 ..
1 MTSSDSKLAG MTSSESIETC EYKMALIEHS LMSRFHAFVN EMIVTKNGRR
51 MEPVLKTSIT GLDPTAMYSV MLDFFVPVDNN RKKYVNGEWI PGCKPEPHVS
101 SCAYIHPDSP NFGSHWMKOP IGFSRVKLTN KATGNPQOIM LNSLHKYEPR
151 IHIMRVGVE SQQVYASHSF QETRFIAVTA YONEDVTSLK IKYNPFAKAF
201 LDGKEERSGN ENYFKDSTKA GSSQNYPRAT TWTPNOSNPT ENQCQYEPGI
251 PPFHYPIPNQ PKTTRRRMS RTQRSHPRKP PTQTQYQDFQ PTNYPILPTD
301 QWSSSIGGHE LDEGHFSLEP VSVDDVTALG FDTPHQAFAP NDLLSIEPSY
351 SLDPYQFTTT WGSRRHMPGY VSNSPIRSLR RPEYFRGYD VSQONYVTMT
401 STDASCVTSS LYETPPSPGVI QRSQEDIQSV YAVTPPSL
11AA_SEQUENCE 1.0
ID Q9GU28 PRELIMINARY; PRT; 334 AA.
AC Q9GU28;
DT 01-MAR-2001 (Tremblrel. 16, Created)
RT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE BRACHYURY PROTEIN.
GN ODT.
OS Olkopleura dioica.
OC Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
OC Olkopleuridae; Olkopleura.
OX NCBI_TaxID=34765;
RN [1]
RP SEQUENCE FROM N.A.
RA Bassham S., Postlethwait J.H.;
RT "Brachyury (T) expression in embryos of a larvacean urochordate,
RT Olkopleura dioica, and the ancestral role of brachyury.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF204208; AAC22592.1; -;
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS50252; TBOX_3; 1.
SQ SEQUENCE 334 AA; 37574 MW; 02AEA233B5800F50 CRC64;
Q9GU28 Length: 334 May 30, 2002 09:32 Type: P Check: 9657 ..
1 MSPLDIKIED APSEADQIEV KLEEPPLMNK FSNLTNEMIV TKNGRMFPV
51 MKVRRKGLDE NAMYTVLDF TAADNHRWKY VNGEWPVGCK PEPQVPSSVY
101 MHPDSPNFGS HWMKQPLSFS KYKLTNKLNO HNGQOIMLHS LHKYEPVHI
151 IKIGGTAGAQ EFVKTQFPFM TRPIAVTAYQ NEEITSLKIR HNPFAKAFLD
201 AEQRKQNDYL PISASANYSN SNGRSAQONN RRRKAENRTQ PYKRPVPVAK


```

251 SESISSIDFG AQNMPSFIQ SMGIPIQAPA RTFSPDGSSA VLGDPMLEPAH
301 ESGSPPLSMP PMNDWCQNGY YSTGLEDPARL ELLV

11AA_SEQUENCE 1.0
ID Q9NHG6 PRELIMINARY; PRT; 110 AA.
AC Q9NHG6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PFEMP1 PROTEIN (FRAGMENT).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=19;
RX MEDLINE=20520970; PubMed=11071291;
RA Taylor H.M., Kyes S.A., Newbold C.I.;
RT "Var gene diversity in Plasmodium falciparum is generated by frequent recombination events.";
RL Mol. Biochem. Parasitol. 110:391-397(2000).
DR EMBL; AF221784; AAF36624.1; -.
FT NON_TER 1 110
SQ SEQUENCE 110 AA; 12828 MW; 65ED0504CF943FA1 CRC64;

Q9NHG6 length: 110 May 30, 2002 09:32 Type: P Check: 7692 ..

1 DLGDIIRGKD LYLGHKLGN KLEARLQTFE QNIKNKNPP LDKLSLEKFR
51 EYWMALNRKE VMKAITCRAG ESDRYSKTIT YGTATSNCK CGHDDQDVOT
101 YLDYVPQFLR

11AA_SEQUENCE 1.0
ID Q9NHG3 PRELIMINARY; PRT; 110 AA.
AC Q9NHG3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PFEMP1 PROTEIN (FRAGMENT).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RX MEDLINE=20520970; PubMed=11071291;
RA Taylor H.M., Kyes S.A., Newbold C.I.;
RT "Var gene diversity in Plasmodium falciparum is generated by frequent recombination events.";
RL Mol. Biochem. Parasitol. 110:391-397(2000).
DR EMBL; AF221827; AAF36667.1; -.
FT NON_TER 1 110
SQ SEQUENCE 110 AA; 13061 MW; 1A30E3C41C607D35 CRC64;

Q9NHG3 length: 110 May 30, 2002 09:32 Type: P Check: 7219 ..

1 DIGDIVRGRD LFLGHKKKN ELEARLQKMF ENIKENNGEL KKLTEQFRE
51 YWMALNRVQV WKAITCAADD NDKYSKITDN GRITFSYDKC GHVNDQVPT
101 NLDYVPQFLR

11AA_SEQUENCE 1.0
ID Q9U218 PRELIMINARY; PRT; 569 AA.
AC Q9U218;
DT 01-MAY-2000 (TREMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Y4IC4A.2 PROTEIN.
GN Y4IC4A.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL032627; CAB63353.1; -.
DR InterPro; IPR000120; Amdase.
DR Pfam; PF01425; Amdase; 2.
SQ SEQUENCE 569 AA; 62918 MW; D0827E77B2A7958D CRC64;

Q9U218 length: 569 May 30, 2002 09:32 Type: P Check: 2401 ..

1 MRKVGKGWHE ADRFOAGARP CLPTMEVLVS EIMSQNNFEK VLIHQKTAPE
51 LQEQDDFETF RRLISSEQL GDSRYLGKCA TELLDLFLR AVYDLDLQSL
101 IDKLOAKDGL NATYVLCAVA RRLDCQTRL DCVAGVMEV GOVTSAPAF
151 QTAQDPTLW YNSEKPPYX GIPFSIKDG ESSLANSPT TLRNIGALP
201 FVTTSISKTP HPRNPWALDC SPGLCGGEA ALVADGAPF GFAPDIDAGS
251 LRIASAFGL VTLKPTRDRF HVSGSNSYGF YTKNQDQVF LKLFIGSAG
301 YRGLEPMSSP APLMDLKLEN KIKIGWFEDD GFNAPVPSNR RAVVDITGIL
351 EKQGEVVKF EMEDIFPPFQ VAQMFETQR PVDNFIPLN YKSNSQCLGR
401 FGKFLNPKL VARISKRAV ISOSSNMKEM CKNLEDIESY KLFIEYKWS
451 LGVDVLICPA FCIPAVPEKY LPELVNTRL TGLFNMDFP AGIVPAGHVT
501 ADDVANLEDE KIFPIDALL RKORDACVNS EAMPNSVQIV GLPNEETVL
551 EVMKIVENLH GPMSNPKGF

```

```

11AA_SEQUENCE 1.0
ID Q9U2X4 PRELIMINARY; PRT; 331 AA.
AC Q9U2X4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Y113G7B.9 PROTEIN.
GN Y113G7B.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Leonard N.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for investigating biology.";
RL Science 282:2012-2018(1998).

```



```

DR EMBL; AL110477; CAB54331.1;
DR InterPro; IPR002892; DUF40.
DR Pfam; PF01838; DUF40; 1.
SQ SEQUENCE 331 AA; 38135 MW; C2F670ACBEDA5462 CRC64;

Q9U2X4 Length: 331 May 30, 2002 09:32 Type: P Check: 4393 ..

1 MEIITQLLCL LGIISAITTI LLNINLVKI VLNPSKRKND MYLFFYRFTL
51 DIFFGAGLFS YIAYTLINME APEFMQYRS LIVLLALPWS HISTCRSIIA
101 LSIISDRSIA TCEPIYYEKN RKKIPNPVL LIGSLGLAE EYMLFGCSY
151 NMEIPKICLV FGCATNOCFF HYWLQORSII FSLIVLFSLI LSIKLLMNS
201 VKHQSNNOI SKANRLALLD TCTVLLFDEL PAFCGHMPT APMEFNNVG
251 SYNPLKITG CAIESTVTVR VLLFRPPENS ASTPVQKLIY TLRPLSLAQ
301 FPIPKNIFE HFLRSRAVTV LYKNIKLIQ F

!!AA_SEQUENCE 1.0
ID Q9N8U9 PRELIMINARY; PRT; 301 AA.
AC Q9N8U9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE POSSIBLE H-SCOL.
GN CHR1.138.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TREU927;
RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
RA Gerrard C., Rajandream M.A., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL359782; CAB95434.1;
DR InterPro; IPR003782; SCOL_Senc.
DR Pfam; PF02630; SCOL_Senc; 1.
SQ SEQUENCE 301 AA; 33578 MW; F03C9365F6CBF559 CRC64;

Q9N8U9 Length: 301 May 30, 2002 09:32 Type: P Check: 7148 ..

1 MSLRRHYITR SISVHNRFS SSGEQKTGGS RGGFFSVLRY GPRNNAELFG
51 LCLGCMCP L SYILINWCEG CSNRNNTOP ELVLGNETEV THGVAAPDKG
101 SKYPLGPF R LRESRTGNYI TDKELEFQDHW TLLYFGFSKC AEVCPSTLRF
151 ITDVMKACDE KLAGDKNLST EAARLQAVFL SVDSRDTPE VLEGFVSKYD
201 PRVRGLTGTS KEIEQAKARAW RVIYSSIDET DEEKSAAREAK GVPWVGADDD
251 TYQLDHSSAI YLVGVGKLIK DFFKENGVA DAVGRLEVHL QDVGFKDTR
301 G

!!AA_SEQUENCE 1.0
ID Q9N8E6 PRELIMINARY; PRT; 136 AA.
AC Q9N8E6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE HYPOTHETICAL 16.2 KDA PROTEIN.
GN CHR1.368.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=TREU927;
RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
RA Gerrard C., Rajandream M.A., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL359782; CAB95597.1;
DR Hypothetical protein.
SQ SEQUENCE 136 AA; 16200 MW; CACDC1CE14964C54 CRC64;

Q9N8E6 Length: 136 May 30, 2002 09:32 Type: P Check: 5512 ..

1 MCQTSTKMSN VEAPASEMKN KERKKEKER EQGKVLKKEK NTKKESFSL
51 FEKKNLFTLL FHEFWFELCL FVSFSEFLS FPRNSVPLVY YIMATVPQPY
101 VTIAITVAIL YIYIYIYIIL KIITFITTST YDYVHC

!!AA_SEQUENCE 1.0
ID Q94216 PRELIMINARY; PRT; 589 AA.
AC Q94216;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 69.4 KDA PROTEIN.
GN F38A5.13.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Fulton B., Stellyes L.;
RT "The sequence of C. elegans cosmid F38A5.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U70854; AAB09157.1;
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR001005; Myb_DNA_bind.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF00249; myb_DNA-binding; 2.
DR SMART; SM00271; DnaJ; 1.
DR SMART; SM00395; SANT; 2.
DR PROSITE; PS50076; DNAJ_2; 1.
DR PROSITE; PS50090; MYB_3; 2.
DR Hypothetical protein.
SQ SEQUENCE 589 AA; 69413 MW; D3GCF39ED6B3EED9 CRC64;

Q94216 Length: 589 May 30, 2002 09:32 Type: P Check: 5664 ..

1 MTGNLQVAI YGFTARSAF EPAGICYETR LIRDKTLGH CTLPIYKNSP
51 VAKPVTEDQR KKSSEPAEL RKDLFDADNE KYEKYLMKLD PNCKNODHY
101 KVLGLSKLRW QATSDEIRFC YRQKVLKHP DKKKHGIWM EKEEYTCIT
151 KAYEQVMSD VKRQAFDSVD HKFNDIIPNE KSINNNNPN ELAPYFOLNS
201 RMSNIKPYPE LKSDATRED VENFYDFWN FQSWREFSYL DEEDKERGED

```

```

251 RYERREMEKQ NKAERERRRK EBAKRIRKLV DIAYAKDPRI IKFKKEQOAK
301 KDKAKEDKOR AIREKQEAID REKREKEEAE AKOKEADARK AKEREREK
351 ERDIAKKAMS OQRKRKLKLA DEAGHWTENP RDKLTEMERI ERICIGFTVD
401 QLRLECEKVE SLISASEIQT ALTDAILKK EAAGAKVTIT EDKNKENEKQ
451 ADKETWTSEE IQLLVKASNT PPPGTVERWV QIADYINEHR KDSTGLPPKT
501 EKOVYKQCKA VQTMNVKLPS TTQNOGLTAL PDEDWSATE QKLEDAIKK
551 HKSSDPERWE KISTEVGTS KKACIRREKY LVQMVKNKK

11AA_SEQUENCE 1.0
ID 046218; PRELIMINARY; PRT; 219 AA.
AC 046218;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE GLUTATHIONE S TRANSFERASE-1.
GN CVGST1.
OS Culicoides varilipennis.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Ceratopogonidae; Culicoides.
OX NCBI_TaxID=46212;
RN [1]
RP SEQUENCE FROM N.A.
RA Abdallah M.A., Pollenz R.S., Tabachnick W.A., Nunamaker R.A.,
RA Murphy K.E.;
RT "Identification of a cDNA clone encoding Culicoides varilipennis
RT Glutathione S-transferase (GST).";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U87958; AAB94639.1; -.
DR HSSP: P30712; 1LJR.
DR InterPro: IPR004046; GST_C.
DR InterPro: IPR004045; GST_N.
DR Pfam: PF00043; GST_C; 1.
DR Pfam: PF02798; GST_N; 1.
DR Transferase.
SQ SEQUENCE 219 AA; 24769 MW; A9EEBD80ACCB1157 CRC64;

046218 Length: 219 May 30, 2002 09:32 Type: P Check: 7893 ..

1 MGIDFYLLPG SSPCRAYQMT AKAVGVDLNL KLTNLMAGEH MKPEFLKLN
51 QHCIPTLVDN GFSLWESRAI QVYLVEKYGK DDSLYPKDVQ QRALVNQRLY
101 FDMGTLYQRF ADYWPQLFA KQAPNPENFK XMEAMGFLN TFLGHRKAYV
151 GDKFTVADLA LAASVATYEV SGFDKPPYPN VQKWFALCKT TLPGYDLNEA
201 GVKNSRIEPL SLNACSGLN

11AA_SEQUENCE 1.0
ID Q17999; PRELIMINARY; PRT; 814 AA.
AC Q17999;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE C15A7.1 PROTEIN.
GN C15A7.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitida; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=99069613; Pubmed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RT Science 282:2012-2018(1998).
RL EMBL: 267735; CAA91530.1; -.
DR InterPro: IPR001175; Neur_channel.
DR Pfam: PF02931; Neur_chan_LBD; 2.
DR PROSITE: PS00236; NEURROT_ION_CHANNEL; UNKNOWN 2.
SQ SEQUENCE 814 AA; 95185 MW; 36464FAB9017B51D CRC64;

Q17999 Length: 814 May 30, 2002 09:32 Type: P Check: 7776 ..

1 MIFPLMILVL FNTIPRVSE ADEESERHQA YLDHRKLKED VLGPLDGSMP
51 SELPVEIGAQ LKLDHISRVD EKTGFLTAL TVKFMWSDER LQWDPSTHSYG
101 FTRTRLEFVD FWLYQIWKPR IYMTNTYERR QSKTIDLISN TLVEMDIHSK
151 GFVMTTTKIL LKTFCLDFEK GYPHDFQNC SFIIPNMNAN DIFLGSSMLK
201 AKFTKDFTDQ QQVVRVQDFQ INDVTEEDLY LFMNRIVLHD LTREPRSLVR
251 SPVSFTIVLQ RKNVYVHLQ YAPMIFICTC LIISGLPSE YALPLLVNLL
301 ISELYLFHNI KDVLPTDFDG TPTIALIATE TLAETMSIIG WKMFITITMT
351 KKOKAYFLNS SKFPLDSGEF NIRVLRRLLT IDKFLMYMLI AQSPGKPTTI
401 FLQLFPPQRA ELRMLTLKVN RQNPFPVLAS QYSVKINISF SLEESSYVNS
451 YFHCARPYPS SCGVTMGNGK KLHGFDDDHG FMSVIVNAEL VSYLVYQPYE
501 SFTFQTWIDE RUKWNPANYS GYREIVEKTF EFRKDNCCWM PIVKYRSYDR
551 RYSELDLLEF SDARTLISYK GEIKTALQTM VTTKQFSFG EYPNDYQNC
601 IMILPNQAD ERFVSPNGY CNPKFENLEH RAVRVHDLHL MGVESNIYYT
651 FANTYFTAEE VTGEFYMAKT EFRFNLMEKR VNKLLNVKLS IPSILISMEL
701 ITAGLPNGY SIFGYAYCFE VEVHGLYVS KILPNDIRGI PYHGALALCF
751 LIETVFLFCW KVFVSYARQK KLFALPNVF ADNRFVLKRF VFYVDRVVCV
801 ALCLQFLSVI YKTR

11AA_SEQUENCE 1.0
ID Q18843; PRELIMINARY; PRT; 660 AA.
AC Q18843;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-JAN-1998 (Tremblrel. 05, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE C54G10.3 PROTEIN.
GN C54G10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitida; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; Pubmed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RT Science 282:2012-2018(1998).
RL EMBL: 275532; CAA99810.1; -.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_GTP_A.

```

DR Pfam: PF00005; ABC_tran: 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
 SQ SEQUENCE 660 AA; 74999 MW; 0A59EB7C8C5079CA CRC64;

Q18843 Length: 660 May 30, 2002 09:32 Type: P Check: 7424 ..

1 MARLALDFLF FRRFRWRLKI LFPCHRSIT AWLAGALLIA TALDQVMTYL
 51 VGIQPSLFYV ALGNKDADTF KILCARGAI ILGKFTLAV FRYLTNMLAI
 101 KSRQVCNLTM HRLYFKROAF EKLSSSGDML DNPDRQLTOD IEKATRILSN
 151 DLAPIATAP FIGYTTWLT YESSGWIGPA AIITYFCIQI IINKMILSPI
 201 VQKVSQEKM EGDFRQRHME VRSNVEAIAF YRAGVLENIM TNOKLKNLIE
 251 TQKSLTEWRM VLNSTINVED YFGILSYLI IGVPEFITHL YDDVSPAELN
 301 GIVSRNAFFY LYLYSFSTV LKLTGDFGEL AGVTHRVML HEELNRLHSD
 351 CLETRDRPST VPSSVVIVAS DEDDKSASRH MQEIHGKQMS LERDEQEEEE
 401 AQYLLGKGTG QEDDMPDDGV AIYVDSATLS PRNDHSHLIV QLLSLQIIQG
 451 QTLITGDSG CGKSSLLRMF AGLWHCSSGK MDCHWRRLTS NLEFLAQKPY
 501 FPGSNTTLRQ QIYYPKALQ VDKVARITQ ILEWVKMEHL VERCGGLDTP
 551 VEMDMKTLIS PGLQRLSLA RVFYTKPRIV FLDESTSAIG FELEMAIYRK
 601 LQEKITFVS IGHRYSLKQF HDMELRVKGR SGWLSLHDID TASIASRTAS
 651 FLGADTVLSM

!!AA_SEQUENCE 1.0
 ID Q9XUC9 PRELIMINARY; PRT; 170 AA.

AC Q9XUC9;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE M162.2 PROTEIN.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Basham V.M.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;
 RA none;

RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology.";
 RL Science 282:2012-2018(1998).

DR EMBL; 282278; CAB05254.1; -;
 DR InterPro; IPR001304; lectin_c.
 DR Pfam; PF00059; lectin_c; 1.

DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 SQ SEQUENCE 170 AA; 19234 MW; 1EB4956690519F17 CRC64;

Q9XUC9 Length: 170 May 30, 2002 09:32 Type: P Check: 6239 ..

1 MFKLLFLAI LIPSSLTDCP DIDGQTRERIK GICYKFTVSL MKYEDARDWC
 51 HHNPAGPSW LAYVPDQETS NFLAVYAGSI FGEGRKHFMI GLNRDPISKI
 101 LSWDTGLSVS YTFNGSNVAQ NYFSENITNT KWNLTGDDDEV HNFVCSYRPS
 151 TVPATVTRQP QARRLAAMKN

!!AA_SEQUENCE 1.0
 ID O45282 PRELIMINARY; PRT; 705 AA.

AC O45282;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE C31A11.5 PROTEIN.

GN C31A11.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurray A.A.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;
 RA none;

RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; 283218; CAB05687.1; -;
 DR InterPro; IPR003862; UPF0063.
 DR Pfam; PF02712; UPF0063; 1.

SQ SEQUENCE 705 AA; 81206 MW; 03C65ACADD52F9D8 CRC64;

O45282 Length: 705 May 30, 2002 09:32 Type: P Check: 676 ..

1 MKIAPLFEFI VCLPATEVCE NMKEYFLAKA ADQKITGVSE QCANDTETWQ
 51 KSLKNAVELS AECLIEQKCT KEELKTIEDN FYAVEQYDAM GKIPLTGLFQ
 101 LPILFDGSYQ ECERISGKKY ATNYCYVLM PGKNATCHMS DGLPTTFEFR
 151 GAVCMPYSCS EODLPTVYNQ VSDQFTACA AFCSSYPVKK TPAFWGFTSF
 201 MAVMIGIALL ATVIDYLKDA LKKEDEKRED SRIQIILLTF SLWTNAELL
 251 SVKEQKRFI KCLDCIRFLS MLWVYTGHTF SYLTPPDQIE SILPFGGRFW
 301 NHLVMNAFYS VDTFELISGL VVSYLEFKTK LKVSQIKSPI TWILEYVHRY
 351 LRLTPPLMEF LGFFVYGYKY FQGPVSAQL NOONGEVDTC QTYWMKNLIY
 401 INNLMSGDTQ CYGITWYLG A DTQLYLVAPI FLIGLYESFA IGALLTAAT
 451 IGSVITPYIL FSTYDLPADF FGNGDATHFY DMIIYIKPMIR CPPEYVGILV
 501 GYLATYGRK KLRLKMALAV TGNIVAFSLG ALCIFSTYDY DNKVKWSIFS
 551 RATTYNESRL AWSFALSWVI VANHMGWGP IDAFMSHPMW QPFGRLSYCA
 601 YIVHYVVLVM YLMIGDASIH FYSSFOIFMY YAVPTVLSY IFAFWWSCLF
 651 EIPFLKLEKM LIELITGGAR DRNREDIEQ KTLKTKENE LMAVEETVOS
 701 TNEKF

!!AA_SEQUENCE 1.0
 ID O01616 PRELIMINARY; PRT; 533 AA.

AC O01616;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 60.8 KDA PROTEIN T19H12.10 IN CHROMOSOME V.

GN T19H12.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Davidson S.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO FAMILY UPF.
DR EMBL; U97009; AAC69033.1; -.
DR WormBep; T19H12.10; CE13770.
DR InterPro; IPR002213; UDPGT.
DR pfam; PF00201; UDPGT; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 499 519 POTENTIAL.
SQ SEQUENCE 533 AA; 60763 MW; E41F11CFC067124 CRC64;

001616 length: 533 May 30, 2002 09:32 Type: P Check: 9544 ..

1 MRLTFSCLC FLKVNYYAK ILIFNPIGCF SHKVFYSKLA DIADHGQEV
51 TLFQPFHLAL KNLDGIKKNK NIEIINYPD HYDDLKLET QTFPDEWDSQ
101 LMNNPVLMAF MLPRILGGEF KTTIQLIKD KEILKKLKDK KFDVAISETF
151 ELTGMYSHF LGVPCIPILS AVRLDIFNEA FGQSSAFGYL TQQSKLAPD
201 AGFLDRLNDV YRDEFSKMAF RGMAYQYNDV IEKAGHYVP YWKDLVKEAP
251 VYMTNSNPYL DEAVPTTATI VHIGTTINL EKMNHVDALP EYEIILKEK
301 ETTVLISFGS VIRSEMPEN FKAGLIKVFE SLPDVIFIMK YEIDDLFEQK
351 KLPKNVHLKK WVPQPSLAD KRVKLFVTHG GLGSTMEVAY TGPALMVPPI
401 FGDQPMNADM LARHGAIAY DKFDLVGKK LLETVRDLVT NPKYEQKAKE
451 LLDVLTNPPI DPMNLMKHL EFAIKFPNLR SQIPEINQVG PIAHYLDVVI
501 VFLIFVSIIT AYISQIVCR ILSRILSKKY KSD

11AA_SEQUENCE 1.0
ID Q04622 PRELIMINARY; PRT; 326 AA.
AC Q04622;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE U2 SMALL NUCLEAR RIBONUCLEOPROTEIN 40K.
GN U2_40K.
OS Trypanosoma brucei brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=427;
RX MEDLINE=93209224; PubMed=8458336;
RA Cross M., Wieland B., Palfi Z., Guenzl A., Roethlisberger U.,
RA Lahm H.-W., Blindereif A.;
RT "The trans-spliceosomal U2 snRNP protein 40K of Trypanosoma brucei:
RT cloning and analysis of functional domains reveals homology to a
RT mammalian snRNP protein.";
RL EMBL J. 12:1239-1248(1993).
CC -1- FUNCTION: THIS PROTEIN IS ASSOCIATED WITH SN-RNP U2. IT BINDS STEM
CC LOOP IV IN CONJUNCTION WITH OTHER TRYPAOSOMAL PROTEIN(S).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 7.
CC -1- SIMILARITY: TO THE N-TERMINAL HALF OF THE HUMAN U2 SNRNP-A'
CC PROTEIN.
DR EMBL; X69934; CAA49552.1; -.
DR HSSP; P09661; 1A9N.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat;
KW Leucine-repeat.

```

```

FT DOMAIN 21 215 LEUCINE-RICH REPEATS.
FT REPEAT 21 33 LRR 1.
FT REPEAT 34 58 LRR 2.
FT REPEAT 84 101 LRR 3.
FT REPEAT 102 126 LRR 4.
FT REPEAT 127 151 LRR 5.
FT REPEAT 152 178 LRR 6.
FT REPEAT 190 215 LRR 7.
FT MUTAGEN 1 20 MISSING: IN 40K-1; REDUCES U2 BINDING.
FT MUTAGEN 154 326 MISSING: IN 40K-2; ABOLISHES U2 AND
FT MUTAGEN 189 326 GENERAL RNA BINDING.
FT MUTAGEN 215 326 MISSING: IN 40K-3; ABOLISHES U2 AND
FT MUTAGEN 215 326 GENERAL RNA BINDING.
FT MUTAGEN 215 326 MISSING: IN 40K-4; ABOLISHES GENERAL RNA
FT BINDING ACTIVITY.
SQ SEQUENCE 326 AA; 36595 MW; 0B42A08A705C1E6A CRC64;

004622 length: 326 May 30, 2002 09:32 Type: P Check: 3173 ..

1 MRLTDTIRR APQFTNALRQ RELDLRGIGI TVLEHTLTF LNDSPDYVNL
51 SONPLARLEY FPGDSAPLAT AAAQNSSAKP ASRMRLRLQT LVVHRNRLTH
101 VSEATCAAVL PNLRAFVADH NEFRELRDL FLSHWKKLEI LSEHNPIITI
151 SEDNARLRAY VVFLCPTLKL VNYQRYTQVD RQNVETMKRE FVGLVEGMRR
201 LEAKQLLOON SAPTEADASA SESYKKIRKR SRHAREASK NSADTADDEP
251 AASKVETPAT PSVAEAGEE EASNALQAR LEALEEKMAA AETEEELMEL
301 QOELTELETL MKHQASKGT KKTRTS

11AA_SEQUENCE 1.0
ID Q95JPS PRELIMINARY; PRT; 235 AA.
AC Q95JPS;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE HYPOTHETICAL 25.0 KDA PROTEIN.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA Terao K., Sugano S.;
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
RT libraries.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB070136; BAB63081.1; -.
KW Hypothetical protein.
SQ SEQUENCE 235 AA; 24993 MW; D5D728E1149BD3F0 CRC64;

Q95JPS length: 235 May 30, 2002 09:32 Type: P Check: 5683 ..

1 MNMSTATTTH SCVPILVFPS YWASFSCCW TRLVTPMCIL LTPQCEDE
51 DENLYDDPLL LNPPEARSS NSKTTTTLGL VVHAADGVA LGAAASTSQT
101 SVQLIVFAI MLHKAPAFG LVSFIMHAGL ERNRIRKHL VFSLAAPWS
151 MVTYGLSKS SKEALSEVNA TGMAMLSAG TFLYVATVHV LPEVGCIGHS
201 HKPDATGGRG LSRLEVAALV LGCLIPILIS VGHQH

11AA_SEQUENCE 1.0
ID Q9TGJ4 PRELIMINARY; PRT; 161 AA.
AC Q9TGJ4;
DT 01-MAY-2000 (Tremblrel. 13, Created)

```

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 4 (FRAGMENT).
GN ND4.
OS Taenia crassiceps.
OG Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC Cyclophyllidae; Taeniidae; Taenia.
OX NCBI_TaxID=6207;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21036605; PubMed-11163447;
RA Nakao M., Sako Y., Yokoyama N., Fukunaga M., Ito A.;
RT "Mitochondrial genetic code in cestodes."
RL Mol. Biochem. Parasitol. 111:415-424(2000).
DR EMBL; AB031287; BAA83540.3; -.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 161 AA; 18553 MW; 4FE537849D6CBDE8 CRC64;
SQ

Q9TGU4 Length: 161 May 30, 2002 09:32 Type: P Check: 134 ..

1 AFLSLSHIV PFLCLFISDW SLSFSFFYC MGHGSLAGIV FCLLMCFYDI
51 SNTRMNVLIK SSINGINLIM IVFSLSLSC SFPPTIQFC EVNLVINSSN
101 LLLLIIFWFC YLFLGLVPL VLCGHLLIRS EFVESGVSC YSYFELVFS
151 CFWCYLGFFI F

!!AA_SEQUENCE 1.0
ID Q9B8W7 PRELIMINARY; PRT; 417 AA.

AC Q9B8W7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4 (EC 1.6.5.3).
GN NAD4.
OS Taenia crassiceps.
OG Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC Cyclophyllidae; Taeniidae; Taenia.
OX NCBI_TaxID=6207;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AMERICAN;
RX MEDLINE=20349913; PubMed-10889225;
RA Le T.H., Blair D., Agatsuma T., Humair P.F., Campbell N.J.,
RA Iwagami M., Littlewood D.T., Peacock B., Johnston D.A., Bartley J.,
RA Rollinson D., Herniou E.A., Zarlenga D.S., McManus D.P.;
RT "Phylogenies inferred from mitochondrial gene orders-a cautionary tale
RT from the parasitic flatworms."
RL Mol. Biol. Evol. 17:1123-1125(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AMERICAN;
RA Blair D., Le T.H., Littlewood D.T.J., Zarlenga D.S., McManus D.P.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR EMBL; AF216699; AAG13172.2; -.
DR InterPro; IPR003918; NADHub_oxdrdctse4.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; Oxidored_q1; 1.
DR PRINTS; PR01437; NUOXDRDTASE4.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 417 AA; 48714 MW; 521EF225852326BC CRC64;

Q9B8W7 Length: 417 May 30, 2002 09:32 Type: P Check: 7748 ..

1 MFIFFSVSWF IVLVLEFVLL LVFSCGVNCF SMINSVYVNG LFFVDSIFFY
51 LFLVLEFGL YSNVMFYNLL SFDTRYLVLV SLLETSLSFC INHCYFFWCF

101 YELSMPLPLY LIFCDHSYSE RFLAGWYFCS YLLTSLPLI LILLYLSYVN
151 NSFYFSSWYA GDDVYLIYY LLSFEFTKV PLVPFHTWLP IYNAEATSIY
201 SIFLSGYIMK LGLLGYVRCs CFIFNVSELM YLFICILCV FFLIVSCSEL
251 DGRWLAFLS LSHIVPELC LFISDMSSLS FSFFYCMGHG LSAIVFCLL
301 WCFYDISNTR NWVLKSSIN GINLIMIVF SLISCSFPT TIQFCEVNL
351 VINSNLLL IIFWECYFL GGLVPLVLCG HLIRSEYE SVGVCSCSYF
401 YFLVSCFWC YLGFFIF

!!AA_SEQUENCE 1.0
ID Q94TE9 PRELIMINARY; PRT; 116 AA.

AC Q94TE9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 3.
GN ND3.
OS Chlorophthalmus agassizi (shortnose greeneye).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Aulopiformes; Chlorophthalmoidae; Chlorophthalmidae; Chlorophthalmus.
OX NCBI_TaxID=143313;
RN [1]
RP SEQUENCE FROM N.A.
RA Miya M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21519020; PubMed-11606696;
RA Miya M., Kawaguchi A., Nishida M.;
RT "Mitogenomic exploration of higher teleostean phylogenies: A case
RT study for moderate-scale evolutionary genomics with 38 newly
RT determined complete mitochondrial DNA sequences."
RL Mol. Biol. Evol. 18:1993-2009(2001).
DR EMBL; AP002918; BAB70021.1; -.
KW Mitochondrion.
SQ SEQUENCE 116 AA; 12978 MW; 10657FE65DFA5AB CRC64;

Q94TE9 Length: 116 May 30, 2002 09:32 Type: P Check: 2807 ..

1 MNLVSTVIMI ALALSSILAF VSFWLPQINP DTEKISPYEC GFDPGASARL
51 PFSLRFFLYA ILFLFDLEI ALLPLPWGD QLTPTPTHTFY WATVALLLT
101 LGLAYEWTQG GLEWAE

!!AA_SEQUENCE 1.0
ID Q94T60 PRELIMINARY; PRT; 116 AA.

AC Q94T60;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 3.
GN ND3.
OS Trachipterus trachipterus (ribbon fish).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Lampridiformes; Trachipteridae; Trachipterus.
OX NCBI_TaxID=143894;
RN [1]
RP SEQUENCE FROM N.A.
RA Miya M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21519020; PubMed=11606696;

```

RA Miya M., Kawaguchi A., Nishida M.;
RT "Mitogenic exploration of higher teleostean phylogenies: A case
RT study for moderate-scale evolutionary genomics with 38 newly
RT determined complete mitochondrial DNA sequences.";
RL Mol. Biol. Evol. 18:1993-2009(2001).
DR EMBL; AP002925; BAB70112.1; -.
KW Mitochondrion.
SQ SEQUENCE 116 AA; 13135 MW; 020ED948A2E92E93 CRC64;

Q94T60 Length: 116 May 30, 2002 09:32 Type: P Check: 3118 ..

1 MNLVVTSLLI ALLSTVLIF ISFWIPQLSP DHEKLSPYEC GFDPLGSARL
51 PFSLRFFLVA ILFLFDLEI ALLPLPWGN HLASPLETFT WASLVLLILT
101 LGLIYEWYQG GLEWAE

11AA_SEQUENCE 1.0
ID Q94T47 PRELIMINARY; PRT; 116 AA.
AC Q94T47;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 3.
GN ND3.
OS Zu cristatus (scaloped ribbonfish).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Lampridiformes; Trachipteridae; Zu.
OX NCBI_TaxID=143325;
[1]
RN SEQUENCE FROM N.A.
RA Miya M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21519020; PubMed=11606696;
RA Miya M., Kawaguchi A., Nishida M.;
RT "Mitogenic exploration of higher teleostean phylogenies: A case
RT study for moderate-scale evolutionary genomics with 38 newly
RT determined complete mitochondrial DNA sequences.";
RL Mol. Biol. Evol. 18:1993-2009(2001).
DR EMBL; AP002926; BAB70125.1; -.
KW Mitochondrion.
SQ SEQUENCE 116 AA; 13145 MW; E8E24CB7AAE295F9 CRC64;

Q94T47 Length: 116 May 30, 2002 09:32 Type: P Check: 2171 ..

1 MNLVMTSLVI ALLSVLIF ISFWIPQLSP DHEKLSPYEC GFDPLGSARL
51 PFSLRFFLVA ILFLFDLEI ALLPLPWGN HLACPQETFI WATLVLLILT
101 LGLIYWSQG GLEWAE

11AA_SEQUENCE 1.0
ID Q94S17 PRELIMINARY; PRT; 116 AA.
AC Q94S17;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 3.
GN ND3.
OS Antigonla capros (deepbody boarfish).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Zeiformes; Caproidae; Antigonla.
OX NCBI_TaxID=143339;
[1]
RN SEQUENCE FROM N.A.
RA Miya M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

```

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21519020; PubMed=11606696;
RA Miya M., Kawaguchi A., Nishida M.;
RT "Mitogenic exploration of higher teleostean phylogenies: A case
RT study for moderate-scale evolutionary genomics with 38 newly
RT determined complete mitochondrial DNA sequences.";
RL Mol. Biol. Evol. 18:1993-2009(2001).
DR EMBL; AP002943; BAB70346.1; -.
KW Mitochondrion.
SQ SEQUENCE 116 AA; 13055 MW; 2A471321E3C3291B CRC64;

Q94S17 Length: 116 May 30, 2002 09:32 Type: P Check: 2075 ..

1 MNLVTIITAI SITLSTILAF VSFWLPQMP DHEKLSPYEC GFDPLGTARL
51 PFSLRFFLVA ILFLFDLEI ALLPLPWGD QLSSPLTFTL WAAAVLTLLT
101 LGLIYEWLQG GLEWAE

11AA_SEQUENCE 1.0
ID Q94SD5 PRELIMINARY; PRT; 116 AA.
AC Q94SD5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 3.
GN ND3.
OS Dactyloptena peterseni (starry flying gurnard).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Scorpaeniformes;
OC Dactylopteroidae; Dactylopteridae; Dactyloptena.
OX NCBI_TaxID=143342;
[1]
RN SEQUENCE FROM N.A.
RA Miya M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21519020; PubMed=11606696;
RA Miya M., Kawaguchi A., Nishida M.;
RT "Mitogenic exploration of higher teleostean phylogenies: A case
RT study for moderate-scale evolutionary genomics with 38 newly
RT determined complete mitochondrial DNA sequences.";
RL Mol. Biol. Evol. 18:1993-2009(2001).
DR EMBL; AP002947; BAB70398.1; -.
KW Mitochondrion.
SQ SEQUENCE 116 AA; 13092 MW; 40C80DF75E078634 CRC64;

Q94SD5 Length: 116 May 30, 2002 09:32 Type: P Check: 3914 ..

1 MNLITTIMI SALLTIVLAF VSFWLPQMP DHEKLSPYEC GFDPLGTARL
51 PFSLRFFLVA ILFLFDLEI ALLPLPWGD QLTSPVTFTL WASSVLLILT
101 LGLIYEWLQG GLEWAE

11AA_SEQUENCE 1.0
ID Q96303 PRELIMINARY; PRT; 534 AA.
AC Q96303;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN PT2 OR PHT4 OR TTF6.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN

```


RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=97045083; PubMed=8927627;
RA Muchhal U.S., Pardo J.M., Raghothama K.G.;
RT "Phosphate transporters from the higher plant Arabidopsis thaliana.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10519-10523(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG;
RA Mdeguie-A-Mdeguie D., Gomez R.-M., Fils-Lycaon B.;
RL Plant Physiol. 115:1288-1288(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG;
RA Mukatira U.T., Muchhal U.S., Raghothama K.G.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=99087485; PubMed=9872450;
RA Okumura S., Mitsuoka N., Shirano Y., Shibata D.;
RT "Phosphate transporter gene family of Arabidopsis thaliana.";
RL DNA Res. 5:261-269(1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
Shen M., Rensing C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC T7F6 genomic sequence.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; U62331; AAB17266.1; -
DR EMBL; AF022872; AAB88291.1; -
DR EMBL; AB016166; BAA34398.1; -
DR EMBL; AC005770; AAC79607.1; -
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SO SEQUENCE 534 AA; 58599 MW; 8186DBBCB8379E CRC64;
Q96303 Length: 534 May 30, 2002 09:32 Type: P Check: 605 ..
1 MAREQLQVLN ALDVAKTQWY HFTAILIAGM GFTDAYDLF CISTVTKLLG
51 RIYYHVEGAQ KPGTLPPNVA AAVNGVAFCG TLAGQLFFGM LGDKLGRKKV
101 YGWTLMWVL CSIASGLSFG HEPRKAVMATL CFFRWLGFSG IGDYPLSAT
151 IMSEYANKKT RGAFVSAVFA MQFGIMAGG IFAILISSAF EAKPPSPAYA
201 DDALGSTIPQ ADLVWRILM AGAIPAMTY YSRKMPETA RYALVAKDA
251 KQAASDMSKV LQVEIEBQQ KLEISKEKS KAFGLSKEF MSRHGLHLG
301 TTSTWFLDI AFYSQNLFOK DIFSAIGWIP PAQSMNAIQE VFRIARAOTL
351 IALCSTVPGY WFTVAFIDVI GRFAIQMGF FFMVFMFAL AIPYHMTHK
401 ENRIGFVIMY SLTFFPANGF PNATTFVVP EIFPARFRST CHGISAASGK
451 LGAMVGAFGF LYLAQNDPKD KTDAGYPPGI GVRNSLIVLG VVNFGLILFT
501 FLVPESKGS LEEMSGENED NENSNDSTRT VPIV
11AA_SEQUENCE 1.0
ID 050040 PRELIMINARY; PRT; 542 AA.
AC 050040;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PUTATIVE PHOSPHATE TRANSPORTER.
GN PHT5 OR T21L14.23.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Okumura S., Mitsuoka N., Shirano Y., Sato S., Shinozaki K.,
RA Shibata D.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AB000093; BAA24280.1; -
DR EMBL; AC003033; AAB91982.1; -
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SO SEQUENCE 542 AA; 59213 MW; 9B63C154899A8AC1 CRC64;
050040 Length: 542 May 30, 2002 09:32 Type: P Check: 992 ..
1 MAKKGKEVLN ALDAKTQWY HFTAILIAGM GFTDAYDLF SIVTKLLG
51 RIYYHVDSSK KPGTLPPNVA AAVNGVAFCG TLAGQLFFGM LGDKLGRKKV
101 YGITLMWVL CSLGSLSGSFG HSANGVAMTL CFFRWLGFSG IGDYPLSAT
151 IMSEYANKKT RGAFIAVFA MQFGILAGG IVSLIVSSTF DHAFRAPTYE
201 VDPVGSIVPQ ADYVWRIVLM FGAIPALITY YWRMKMPETA RYALVARNT
251 KQAASDMSKV LQVDLIAEE AQSNSSNSNP NTFGLFTRE FARRHGLHL
301 GTTTFWFLD IAYSSNLFQ KDITYAIGWI PAAETMNAIH EVFTVSAQT
351 LIALGTVPG WFTVAFIDI LGRFFIQMG FIFMTIFMFA LAIPYDHRH
401 RENRIGFLIM YSLTMFANF GPNATTFVVP AEIFPARLRS TCHGISAASG
451 KAGATVGAFG FLYAQSDDS EKTDAGYPPG IGVNSLLML ACVNFGLIVF
501 TLLVPESKCK SLEISREDE EQSGDITYVE MIVANGSRKV PV
11AA_SEQUENCE 1.0
ID 096302 PRELIMINARY; PRT; 524 AA.
AC 096302;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN PT1 OR PHT1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=97045083; PubMed=8927627;
RA Muchhal U.S., Pardo J.M., Raghothama K.G.;
RT "Phosphate transporters from the higher plant Arabidopsis thaliana.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10519-10523(1996).
RN [2]
RP SEQUENCE FROM N.A.


```

RC STRAIN=COLUMBIA;
RX MEDLINE=9738151; PubMed=9192698;
RA Mitsuoka N., Okumura S., Shirano Y., Sato S., Kato T., Harashima S.,
RA Shibata D.;
RT "Overexpression of an Arabidopsis thaliana high-affinity phosphate
RT transporter gene in tobacco cultured cells enhances cell growth under
RT phosphate-limited conditions.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:7098-7102(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; U62330; AAB17265.1; -.
DR EMBL; D86608; BAA21504.1; -.
DR EMBL; D86591; BAA21503.1; -.
DR EMBL; AB025638; BAA97414.1; -.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 524 AA; 57616 MW; 2D12D50987127BD6 CRC64;

Q96302 Length: 524 May 30, 2002 09:32 Type: P Check: 8746 ..

1 MAEQQLGVLK ALDVAKTQLY HFTAIVIAGM GFTDAYDLF CVSLVTKLLG
51 RIYFNPESA KPGSLPPHVA AAVNGVALCG TISGQLFFGW LGDKLGRKKV
101 YGLTLVMIL CSVASGLSFG HEAKGVMTTL CEFREWLFGG IGGDYPLSAT
151 IMSEYANKKT RGAFTAAVFA MGVGILAGG FYALAVSSIF DKKFPAPTYA
201 VNRLSTPPQ VDIYIRIIVM FGALPALTY YWRMKMPETA RTALVAKNI
251 KQATADMSKV LQTDIELEER VEDDVKDPKQ NYGLFSKEFL RRHGLHLGTT
301 TSTWFLLDIA FYSQNLFOKD IFSAGIWPK AATMNATHEV FRIARAQTLI
351 ALCTVPGYW FTVAFIDTIG RERIQNGFF MMTVFMAFA EPPNHWIKPE
401 NRIGFVVMYS LTFEEANFGP NATTFIVPAE IFPARLRSTC HGISAAGKA
451 GAIVGAFGFL YAAQSDQKAK VDAGYPPGIG VKNSLIMLV LNFIGMLFTF
501 LVPEPKGKSL EELSGEAEVS HDEK

!!AA_SEQUENCE 1.0
ID 048640 PRELIMINARY; PRT; 524 AA.
AC 048640;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER.
GN PHT2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Mitsuoka N., Okumura S., Shibata D.;
RL Soil Sci. Plant Nutrition 43:971-974(1997).
RN [2]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AB000094; BAA24282.1; -.
DR EMBL; AB025638; BAA97416.1; -.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 524 AA; 57644 MW; 54664D4CD3ABE9D7 CRC64;

O48640 Length: 524 May 30, 2002 09:32 Type: P Check: 8027 ..

1 MAEQQLGVLK ALDVAKTQLY HFTAIVIAGM GFTDAYDLF CVSLVTKLLG
51 RIYFNPESA KPGSLPPHVA AAVNGVALCG TISGQLFFGW LGDKLGRKKV
101 YGLTLVMIL CSVASGLSFG NEAKGVMTTL CEFREWLFGG IGGDYPLSAT
151 IMSEYANKKT RGAFTAAVFA MGVGILAGG FYALAVSSIF DKKFPAPTYA
201 VNRLSTPPQ VDIYIRIIVM FGALPALTY YWRMKMPETA RTALVAKNI
251 KQATADMSKV LQTDIELEER VEDDVKDPKQ NYGLFSKEFL RRHGLHLGTT
301 TSTWFLLDIA FYSQNLFOKD IFSAGIWPK AATMNATHEV FRIARAQTLI
351 ALCTVPGYW FTVAFIDTIG RERIQNGFF MMTVFMAFA EPPNHWIKPE
401 NRIGFVVMYS LTFEEANFGP NATTFIVPAE IFPARLRSTC HGISAAGKA
451 GAIVGAFGFL YAAQSDQKAK VDAGYPPGIG VKNSLIVLV LNFIGMLFTF
501 LVPEPKGKSL EELSGEAEVS HDEK

!!AA_SEQUENCE 1.0
ID 048639 PRELIMINARY; PRT; 521 AA.
AC 048639;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER.
GN PHT3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Mitsuoka N., Okumura S., Shibata D.;
RL Soil Sci. Plant Nutrition 43:971-974(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AB000094; BAA24281.1; -.
DR EMBL; AB025638; BAA97415.1; -.

```

DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 521 AA; 57257 MW; 5EB3FC6D42590E2F CRC64;

048639 Length: 521 May 30, 2002 09:32 Type: P Check: 4821 ..

1 MADQQLGVLK ALDVAKTQLY HETAIVIAGM GFEYDAYDLF CVALVTKLGG
51 RLYYFNPTSA KPGSLPPHVA AAVNGVALCG TLAGQLFFGW LGDKLGRKKV
101 YGITLIMIL CSVASGLSLG NSAKGVMTTL CEFREWLFGG IGDYPLSAT
151 IMSEYANKKT RGAFLAAVEA MQVGILAGG FVALAVSSIF DKKEPSPTYE
201 ODRFLSTPPQ ADYIWRIVM FGALPALTY YWRMKPETA RYTAIVAKNI
251 KQATADMSKV LQTDLELEER VEDVYKDPKK NYGLFSKEFL RRHGLHLIGT
301 TSTWFLDIA FYSQNLFOKD IFSALGWIPK AATMNAIHEV FKARAQTLI
351 ALGSTVPGYW FTVAFLDIIG RFAIQLMGEF MMTVEMFAIA FPNHWILPD
401 NRIGFVWYS LTFEFANFGP NATFTVPAE IFPARLRSTC HGISAATGKA
451 GAIVGAFGFL YAAQPODKTK TDAGYPPGIG VKNSLIMLGV INFVGMLETF
501 LVPEPKGKSL EELSGAEVD K

11AA_SEQUENCE 1.0
ID Q9ZQX5 PRELIMINARY; PRT: 3449 AA.

AC Q9ZQX5;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE TRANSPORT PROTEIN.
GN T10P11.5 OR AT4G02660.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Kaplan N., Johnson D., Schutz K., Gnoj L., Hoffman J., Till S.,
RA de la Bastide M., Granat S., Hameed A., Gottesman T., Hasegawa A.,
RA Shohdy N., Barnell L., Dedhia N., Johnson A.F., Lodhi M.,
RA Martienssen R., Chen E.Y., Wilson R., McCombie W.R.;
RT "Sequence of A. thaliana BAC T10P11 from chromosome IV.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vill D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 3 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AC002330; AAC78268.1; -.
DR EMBL; AL161495; CAB77751.1; -.
DR InterPro: IPR000409; Belge_BEACH.
DR InterPro: IPR001680; WD40.
DR Pfam; PF02138; Beach; 1.
DR Pfam; PF00400; WD40; 3.
DR ProDom; PD007848; Belge_BEACH; 1.
DR SMART; SM00320; WD40; 2.
DR PROSITE; PS50197; BEACH; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.

DR PROSITE; PS50082; WD_REPEATS_2; 1.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 3449 AA; 385006 MW; 6BB0DBA778690184 CRC64;

Q9ZQX5 Length: 3449 May 30, 2002 09:32 Type: P Check: 8548 ..

1 MKWGTLLKDL KDKVGAETT ADLAGEAIS DPTTPSSSQ ASPSSFAAL
51 AQHDFNLSP TSRDKLIEL DFKRYWEEER SSSSEOEKEA ALNLSVNTFC
101 RLVKQHAYND QLVTMLVEPH IFSEYIGRAF VADVEKLKDQ SSHGANLLTA
151 IEVLASGPPD KQSLDSGIL CCLHTFNAF LTVSAVEGE KTVNYEKEVE
201 GSVVNIMKAL ASHPAAQSL IEDDSLQLLF KMWANGSLMA FSRRKVLVS
251 FHNQLHKNA MQILGLLVN DNGSTASYR KHHLLKVLMM AVKDEPDPCG
301 DSAYTVGIYD LLECEVELSY RPEYGVRLK DDIRNAHYH FLYQFALILS
351 SMPKDIYFAF DHSSPHKRG SSDKQOPPL SLKTRONDS EKQOSLSLNS
401 RQDEFALKH ESPALSRLLD VLVTLAQGP IESSGTSTSL LSQTFKLGYG
451 RROTPSANNR YDEPCEGSG KVKDLEAVQM LQDIFLKAEN KDQLAEVLNR
501 MKIFTSHLE NYRICOELKT VPLLVLNMG FPSLQELIL KILEYAVTVV
551 NCPVEQELLS LCFLLQPID SELKHTILSF FVKLSFDQ YKKVGEVGV
601 LEVLQDDLKQ HKLLRGPQY SGVSNHLDRV PSSPFRQHL DSQDAITSSP
651 KLMESGSKL PIFEVERTT VGWDCMISLL KNSQVQGEAF RSANGVTVIL
701 PELADEHRT SILRIFSCLI TGDIKQVHE ELEALIDVLK SGWYTRVSGD
751 QYKLHYEVRD DIMGALWRIY GVNSAQRF GEATGFSLLL TTIHTFGGEE
801 ECRDESHLVY YIKLFKHLR LITTVACENA INRMKLSVI TSQTFYDLLV
851 ESGLICVDLE RHVIQLLEL ALEVLVPPFL TSESMAAEM AECEKASFLV
901 KTASGFNPD KOKIYNAGV RVLIRSLLC TPRLQLEFLN LLERLARASP
951 FNKETLTSAG CVELLEIY PFIQSSPFL SHALKIVEVL GAYRLSPSEL
1001 KMLCRYVMQM RVMNSGPSLI GMMKILME EDTGLECVSL APFVENDMSK
1051 TGHASVQVSL GERSWPPAAG YSFVCVQFR NFLTPQELLES EYKAGGSSK
1101 TPILSGQOSE QNIFRIFSVN AISNGSPSYA ELYFQEGIL TLATSNSNL
1151 SFSGLETEEG KWHHLAVVHS KPNALAGLFO ASVAIYIIDG KLRHMGKLG
1201 SPSVPGKSLQ VIIGTSATCA RACGDSMAI LDLLDTDMSS GIQKPEDSNR
1251 QGDSKAHCSSG IWDLDRLGN LSIQLPCKL IFAFDGTCSE FMRATGSFSL
1301 VNLVDPISAA ASLIGGIRF GRLVGNVSLC RQNVIGNSIR PVGMAVYLA
1351 LVEAESRDM LHMALSLIAC ALHONSQNVK DMETITGYHL LALFLRPKMA
1401 LFDMQCLEIF FOISACEAF SEPKLESQ TTISMPTEI IPENNTEDPT
1451 LCKFOYETSS VSGHGDMDF SGRKDSFSL SELEMGDNPV ETSNCIVLSN
1501 ADMEVHVLID WTLWVTAPVS IQIASLGFLE NLISILMYRS HNLAILROIN
1551 LVKHLVTLQ RGDVEVLVLE KLVILLRCIL ENGFLTEPEL DVYRAIMTF
1601 NPPEIKSONS SMRESMKHV IYRNLVLEML IDLYVTIKAE ELLEQWHKTV

1651 SSKLITYFLD GAVHPSSMRW IMTLGVCLT SSPNFSLKFF ASGGYQGLVR
1701 VLQSFYDSDP IYIILFCLIF GKPYVPLPE VRMLDFHALM PDDGSHVELN
1751 FVLLDSVVA MAKSTPRLI MOSMLAHQSG NLSQVSARCV AELVEGYADM
1801 TGELOQKALM HKTAARLMG GEASAPATAT SVIRFMVILA KMCPOFSAC
1851 KNTFELQKA DLYFSCVRAF HAVKLAKOLS MKAEEQNTIG GDDSSVEGNF
1901 CRVSHQDMST KTSISAGSFP QDOTSSVISV DMYPSPDYVA VDKVENELTF
1951 PPGESNKSFO GREYIAKQDG DHVGSVSASS EMKSLDLTGS SSQVQPIDSR
2001 SSESFSMLES PLISEKSSLE VPFIPPSPKS STISTPHPSH ISVSEFDASS
2051 DOSSGSQSS AVHTLTISP KVLLETDESG YGGGPCSAGA SAVLDFMAEV
2101 CADIMTEQIK AVQALESITE MLPLYVDPEC VVVFQGLCLS RVMNYLERRF
2151 LRDEEDDDKK LDKRKWSANL DAFCCMIYDR VYMGAPPOPT GVLRTLEFL
2201 SILOLANKDG RVEEVTSSGK GLLSIGRATR QLDAYVHSIL KNTNRTILYC
2251 FLPSFLITIG EEDLPSRLGL LVESTKKQTS KLSGESGID VSAVLQLIVA
2301 NKNIIICPSN LDTDLNCCIC VNLISLHDQ RKNVQMASN IIKYLLVHRK
2351 SALEDLLVKK PHRGQKFDVL HGGFDRLITG NLPEFSKWLE SSEOITTKVL
2401 EOGAAVMMIQ YIAGSAKFPD VRMGKMDGRR TREMRKRLRD TSKLDLKHWE
2451 QVNERRYALE VVRDAMSAEL RYVRQNKYGL ILHAESVWPT HLQOLVHERG
2501 IFPMRISHGV EDLKWOLCPI EGPYRMKKL ERCKLKIDSL HNLLEGKLEL
2551 GEIELLSKKS EDGLVISDMD SEPAFLISEL YSESFSEAD DLKDVPSARN
2601 GWNNDRATST NAASLHNSLS FGGKSSSTAY SVPISVNTDE KSETGSPKIS
2651 SSGKMEIKH VEESEKELK DDGEYLIRPY LEHLEKIRFR YNCERVVGID
2701 KHDGIFLIGE LCIYVIENFY IDDHGICIEK ECEDELSIID QAQGLKXOFH
2751 GSLESKSKSS TLMSTTIKIG AVGGRWAYG GGAWGKEKYR VTGNLPHPMH
2801 MMKLDVHEI LKRDYELRRV AVEIFSMDCG NDLLVFHKKE REEVEFNLLA
2851 MNLPRNSMLD TTISGSAQOE SKEGSRLEKL MAKSTTKRMQ NGELISNFOYL
2901 MHLNTLAGRG YSDLTOYPVF PWILADYDGE SLDLSDPNNF RKLDKPMGCO
2951 TPGESEEFRR RYESWDDPEV PQFHYGSHYS SAGIVLFYLI RLPFSAENG
3001 KLOGGQFDHA DRLENSIRET WLSAAGKQNT SDVKELIPEF FYMPEFLENR
3051 FNLDLGEKQS GDKVGDVILP PWARGSVREF IRKHREALEB DYVSENLHHW
3101 IDLIFGHKOR GKAAENAVNV FYHYTEGNY DVDAVTDPPAM KASILAQJNH
3151 FGQTPKOLFQ KPHVKRRTDR KVPHPPLKHS MHLVPRNIRK CSSSINOJIT
3201 FNDKLLLTGA NCLLKPRGYK KIIRWGEPDR TLRFMSYDOD KLLSTHENIH
3251 EGNQIQACAV SHDGRIVTG AEDGLVSVMR VSKDGPGRGR RLRLKESICA
3301 HTAKVICLRV SQPYMMIASS SDDCTVIIMD LSSLSEVRQL PNFSVPYTVV
3351 YINDLTGEIV TAAGSVLAWV SINGDCLSYV NTSQLPTDLI VSVAGSTPESD
3401 WLETTWYVTG HQSGALKVMR MVHCTDPKIC CYRMOKOSE TESRRFPDN
11AA_SEQUENCE 1.0

ID Q9ZWT3 PRELIMINARY; PRT; 516 AA.
AC Q9ZWT3;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=99087485; PubMed=9872450;
RA Okumura S., Mitsuoka N., Shirano Y., Shibata D.;
RT "phosphate transporter gene family of Arabidopsis thaliana.";
RL DNA Res. 5:261-269(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty p1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AB005746; BAA34390.1; -;
DR EMBL: AB025638; BAA97413.1; -;
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr.1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Transmembrane.
SQ SEQUENCE 516 AA; 56246 MW; 97B2BFEFA894A145 CRC64;
Q9ZWT3 Length: 516 May 30, 2002 09:32 Type: P Check: 9638 ..
1 MANEEQGSIL KALDVAKTQW YHTAVVVGSG MGEFTDSYDL FVISLITKLL
51 GRIYYQVPGS SSPGLPDGI SAAVSVAFAG GTFIGQIFFG CLGDXLGRKR
101 VYGLTLLIMT ICSICGSLG GRDPKTVMT LCFERFWLGF GIGDYPPLSA
151 TIMSEYSNKR TRGAFFIAVF GMGIGILAA GAVSLVSAY FESKFPSPAY
201 ILDGAASTVP QADYWRIL MVGALPALLT YYWRMKMPET ARYTALVSKN
251 AEQALDMTK VLVNDIEASA AKNDQARVSS DEFGLFSMKF LRRHGLHLIG
301 TASTWFLDDI AFYSQNLFOK DIETTIGWLP SAKTMAIOE LYMIKAQTI
351 IACCSVPGY FFTVGFIIDYM GRKKIQMGF AMMTIFMLSL AIPYHHWTLF
401 ANRIGFVLY SFTFEFSNFG PNATTFIVPA EIFPARIRST CHGISAASGK
451 AGAMVGSFGF SALVKALGMS NTLYIMAGIN LLGLLTFTI PETNGKSLEE
501 LSGETEPEKI KEKIVV
11AA_SEQUENCE 1.0
ID Q9SQY1 PRELIMINARY; PRT; 558 AA.
AC Q9SQY1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE HYPOTHEtical 63.2 KDA PROTEIN.
GN F13M14.22 OR F18K10.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
Maitl R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome 3 BAC F13M14 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F18K10 genomic sequence.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011560; AAC51388.1; -
DR EMBL; AC013428; AAF76351.1; -
DR InterPro: IPR003441; NAM.
DR Pfam: PF02365; NAM; 1.
KW Hypothetical protein.
SQ SEQUENCE 558 AA; 63244 MW; F0BFBCA928C48874 CRC64;
Q9SOY1 Length: 558 May 30, 2002 09:32 Type: P Check: 9617 ..
1 MGRGSVTSLA PGRFRHPTDE ELVRYLLRK ICNKPKFDA ISVTVYKSE
51 PWDLPDKSRL KSRDLEWYFF SMLDKKRRNG SKTNRATEMG YWKTGKDR
101 ILNGSKVVGK KKTLYVHKGR APRGERTNWV MHEYLVDQD LDKTGVDQDA
151 FVLGRIFQKS GSGPRNGEQY GAFVEEEME EEDDMTFVPD QEDLGSEDHV
201 YVHMDIDQK SENEVYVDAL PIPLNFIHG SSNNVETNYS DSINYIQOTG
251 NYMDSGGYFE QPASEYEKDO KPITRDRDGS LQNEGIGCGV QDKHSETLOS
301 SDNIFGTDTs CYNDPVESN YLIGEAFLDP NSNLENDGL YLETNDLSST
351 QQDGFEDFY LTFEDEFDP SQLMGNEDEV FQDELFOEV EYKELEKEET
401 SRSHVVEEK EKDEASCSKQ VDADATEFER DSGISENAD YKYPLLKKAS
451 HMLGAIPAPL ANASEPPTKD AAIRLHAAQS SGSVHTVAGM ITISDSNMGW
501 SYGKNENLDL ILSGLVGNG TAPEKSGNSS ANAMLIEMCF WYLLLSVSFK
551 VSILVSSR
11AA_SEQUENCE 1.0
ID Q9AU01 PRELIMINARY; PRT; 540 AA.
AC Q9AU01;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER 1.
GN PT1.
OS Lupinus albus (White lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genistaeae; Lupinus.
OX NCBI_TaxID=3870;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROTEOID ROOT;
RA Liu J., Unde-Stone C., Li A., Vance C.P., Allan D.L.;
RT "Characterization of phosphate transporters with enhanced expression
in proteoid roots of white lupin.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF305623; AAK38196.1; -
SQ SEQUENCE 540 AA; 59383 MW; 49BBE0B807105D35 CRC64;

Q9AU01 Length: 540 May 30, 2002 09:32 Type: P Check: 5248 ..
1 MAREHIHLN ALDVAKTQWY HSLQIIAGM GFTDAYDLF CISLVTKLG
51 RIYVHDGAG KPGTLPPNVS AAVNGVAFCG TLSQLEFGW LGDKMGRKKV
101 YGMTLMIMVI CSIGSLSEFG HSPNSVLATL CFFRFWLGGF IGGDYPISAT
151 IMSEYNNKKT RGSFIAVFA MQRGILGGG IPALITISAF KARFDAPAYE
201 VDPVGSVTPQ ADYIIRIIVM VGALPALTY YWRMKMPETA RYALVAKNA
251 QQAADMSKV LQVEIQSETN KEAAGKPSF GLFSKEFLRR HGLHLGTAS
301 TWFLIDIAFY SQNLFQKDI F SAIGWIPPAK TMAALDEVYR IARAQTLIAL
351 CSTVPGWFT VALIDRIGRF AIQLNGFFFM TVEMFALAIP YDHWTHKDN
401 IGFVVIYSLT EEFANEGPNA TTFVVPALIE PARFRSTCHG ISSASGKIGA
451 IVGAFGFLYL AQNKDKSKTD AGYPAGIGVK NSLIVLGVN ILGFCFTFLV
501 PEPNGKSLEE MSGENEKEEP TKEGSYNNNN NNNNRTVPYV
11AA_SEQUENCE 1.0
ID Q9AU00 PRELIMINARY; PRT; 539 AA.
AC Q9AU00;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER 2.
GN PT2.
OS Lupinus albus (White lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genistaeae; Lupinus.
OX NCBI_TaxID=3870;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROTEOID ROOT;
RA Liu J., Unde-Stone C., Li A., Vance C.P., Allan D.L.;
RT "Characterization of phosphate transporters with enhanced expression
in proteoid roots of white lupin.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF305624; AAK38197.1; -
SQ SEQUENCE 539 AA; 59135 MW; 36C30F31F34742A6 CRC64;
Q9AU00 Length: 539 May 30, 2002 09:32 Type: P Check: 478 ..
1 MARDELSVLT ALDVAKTQWY HFTALIVAGM GFTDAYDLF SIANTYKLG
51 RIYVHDGAG KPGTLPPNVS LAVNGVALCG TLTGQVEFGW LGDKMGRKKV
101 YGLTLAIMVF ASLASGLSFG HTAKGVSTL CFFRFWLGGF IGGDYPISAT
151 IMSEYNNKKT RGAFLAIVFA MQRGILGGG IYSLVISTIF DHAYSAPAYE
201 VDAEASIVPQ ADYVWRILM LGALPALTY YSRMKMPETA RYALVARN
251 KQAADMSKV LQVEIEAEOD MYDKLSEKDN NNGGLFTKEF AKRHGKHLIG
301 TCTTWELLDI AYYSSNLFQK DIYSSIGWLP PAKEMNAIHE VRFVSKAOVL
351 IALCGTVPGY WFTVAFIDHL GREFIQLMGF FEMTVEMFAL AIPYDHWTKK
401 DNRIGFLVWY ALTFEPANFG PNSTTFVPA EIPPARLRST CHGISSAAGK
451 AGAIVGAFGF LYASQSKDAT KRDVGYPAGI GMRNTLIVLA VCNCLGMFFT
501 FLVPESKGS LLELSEGENED EGTNETSQOP AASNRTVPV
11AA_SEQUENCE 1.0

ID Q9AVR0 PRELIMINARY; PRT; 536 AA.
AC Q9AVR0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN PT1.
OS Sesbania rostrata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Sesbania.
OX NCBI_TaxID=3895;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RA Aono T., Oyaizu H.;
RT "The expression of phosphate transporters in response to phosphate starvation is different from that of H⁺-ATPase or PEPC in Sesbania rostrata."
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AJ286743; CAC28218.1; -
DR SEQUENCE 536 AA; 58625 MW; CCDC7E1584277363 CRC64;
SQ

Q9AVR0 Length: 536 May 30, 2002 09:32 Type: P Check: 6827 ..

1 MAGELGLVNA LDVAKTQWYH FTAIVAGMG FETDAYDLFC ISLVTKLGR
51 IYSDPKAPK PGLPPNVA AVTGVALCGT LAGQLFFGWL GDKLGRKKY
101 GITLLIMVVC SLASGLSFGS SAKGVMATLC FFRFWLGFGI GGDYPLSATI
151 MSEYANKKTR GAFIAVAFAM QGFGIMAGGI VALIVSAGFD HKYKVPSTQE
201 NPQASLVLPQ FDYVWRIVLM FGALPALTY YWRMKMPETA RYALVAKNA
251 QQAATDMSKV LKVELVEVEE KVQKITEADT NKYGLFSKEF AKRHGLHLVG
301 TTTTWFLLDI AFYSQNLFOK DIFSAIGWIP PAKEMNAIHE VYKIRARQTL
351 IALCSTVPGY WFTVAFIDYM GRAFIQLMGF FEMTVEMFAL AIPYDHWTKK
401 ENRIGFVVMY SLTEFFANFG PNATTFVPA EIFPARLRST CHGISAAAGK
451 AGAIVGAFGE LYAAQSKDPT KTDKGYPTGI GIKNSLIMLG VINFAGMLFT
501 LLVPESKGKS LEELSGENED GAEGIELPAS ARTVPV

!!AA_SEQUENCE 1.0
ID Q9AVQ9 PRELIMINARY; PRT; 540 AA.
AC Q9AVQ9;

DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN PT2.
OS Sesbania rostrata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Sesbania.
OX NCBI_TaxID=3895;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RA Aono T., Oyaizu H.;
RT "The expression of phosphate transporters in response to phosphate starvation is different from that of H⁺-ATPase or PEPC in Sesbania rostrata."
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AJ286744; CAC28219.1; -
DR SEQUENCE 540 AA; 58947 MW; 7BA0EA8649BF187A CRC64;
SQ

Q9AVQ9 Length: 540 May 30, 2002 09:32 Type: P Check: 3365 ..

1 MAGGGGQLGV LNALDVAKTQ WYHFTAIVIA GMGEFTDAVD LFCISLVTKL
51 LGRIYYTDYN KRPKGVLPNN VQAAVTGVAL CCTLGQLFF GWLGDKLGK
101 KYYGITLMM VVCSIASGLS FGSSAKGVMA TLCEFRFWLG FGIGDYPIS
151 ATIMSEYANK KTRGAFIAAV FAMQFGIMA GGIIVALIVSS AFDHKKYVPA
201 YQEDAKASMV LPAFDYVWRI ILMGAVPAA LTYWRMKMP ETARYTALVA
251 KNAKQAAADM SKVLQVEIEA EEDKVQHLIE SQNQTFGLFS KEFAKRHGLH
301 LVGTTTWFL LDIAFYSQNL FQKDIPTAIG WIPPAKEMNA IHEVYRIARA
351 QTLIALCSTV PGYWFTVAFI DYMGRAIQL MGEFFMTVEM FALAIPYHHW
401 TLKENRIGFV VMXSLTFFFA NFGPNATTFV VPAEIFPARL RSTCHGISAA
451 AGKAGAIVGA FGFLYAAQST DPAKTDAGYP TGIGVKNSLI MLGVINFRGI
501 IFTLLVPESK GKSLEELSGE TEEDGVEAIE AAASSRTVPV

!!AA_SEQUENCE 1.0
ID Q96243 PRELIMINARY; PRT; 524 AA.
AC Q96243;

DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN APT1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RX MEDLINE=97177798; PubMed=9025304;
RA Smith F.W., Ealing P.M., Dong B., Delhaize E.;
RT "The cloning of two Arabidopsis genes belonging to a phosphate transporter family."
RL Plant J. 11:83-92(1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: Y07681; CAA68945.1; -
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR Transmembrane.
KW SEQUENCE 524 AA; 57610 MW; 62BD256C59834B5D CRC64;
SQ

Q96243 Length: 524 May 30, 2002 09:32 Type: P Check: 8141 ..

1 MAEQQLGVLK ALDVAKTQLY HFTAIVIAM GFTDAYDLF CVSLVTKLIG
51 RIYYNPESA KPGSLPPHVA AAVNGVALCG TISGQLFFGW LGDKLGRKKV
101 YGTLFIMIL CSVASGLSFG NEAKVMTTL CFFRFWLFGG IGDYPLSAT
151 IMSEYANKKT RGAFIAAVEA MQGVGILAGG EVALAVSSIF DKKEPAPTYA
201 VNRAALSTPPQ VDIWRIIVM FGALPALTY YWRMKMPETA RYALVAKNI
251 KQATADMSKV LQTDIELEER VEDVKDPPRQ NYGLFSKEEL RRHGLHLGCT
301 TSTWFLDIA FYSQNLFOKD IFSAGIWPK AATMNATHEV FRIARAQTLI
351 ALCSTVPGYW FTVAFIDTIG RFKIQLNGFF MMTVFMFAIA FPNHWTKPE
401 NRIGFVVMYS LTFEFANLGP NATTFIVPAE IEPARLRSTC HGISAAGKA
451 GAITGAFGL YAAQNDKAK VDAGYPPGIG VKNSLIVLGV LNFIGMLFTF

```

501 LVPEPKGKSL EELSGEAEVS HDEK

11AA_SEQUENCE 1.0
ID Q96264 PRELIMINARY; PRT; 524 AA.
AC Q96264;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN APT2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RX MEDLINE=97177798; PubMed=9025304;
RA Smith F.W., Ealing P.M., Dong B., Delhalze E.;
RT "The cloning of two Arabidopsis genes belonging to a phosphate
RT transporter family."
RL Plant J. 11:83-92(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; Y07682; CAA68946.1; -.
DR InterPro: IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 524 AA; 57642 MW; CC12C41887127BCB CRC64;

Q96264 Length: 524 May 30, 2002 09:32 Type: P Check: 8941 ..

1 MAEQQLGVLK ALDVAKTQLY HETAIVIAM GFETDAYDLF CVSLVTKLIG
51 RIYYNPESA KPGSLPPHYA AAVNGVALCG TLSCQLFFGW LGDKLGRKKV
101 YGLFLVMIL CSVASGLSFG HEAKGVMTL CFFRWLFGF IGDYPLSAT
151 IMSEYANKKT RGAFIAVFA MGVGILAGG EYALAVSSIF DKRPAPTYA
201 VNRAIETPQ VDYIWRIVM FGALPALTY YWRMKMPETA RYALVAKNI
251 KQATADMSKV LQTDIELEER VEDDVKDPKQ NYGLFSKEFL RRRGLHLIGT
301 TSTWFLDIA FYSQNLFOKD IFSALGWIPK AATMNATHEV FRIARAQTLI
351 ALCSIVPGYW FTVAFIDTIG RFKIQLNGFF MMTVEMFAIA FPNHWIKPE
401 NRIGFVVMYS LTFEAFNFGP NATTEIVPAE IFPARLRSTC HGISAAGKA
451 GAIVGAFGFL YAAQSODKPK VDAGYPPGIG VKNSLIMLGV LNFIGMLTFE
501 LVPEPKGKSL EELSGEAEVS HDEK

11AA_SEQUENCE 1.0
ID Q24029 PRELIMINARY; PRT; 538 AA.
AC Q24029;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN LEPT1.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MONEYMAKER, TISSUE=ROOT HAIR;
RX MEDLINE=98408214; PubMed=9737001;
RA Darim P., Brunner S., Persson B.L., Amrhein N., Bucher M.;

```

```

RT "functional analysis and cell-specific expression of a phosphate
RT transporter from tomato.";
RL Planta 206:225-233(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; Y14214; CAA74607.1; -.
DR InterPro: IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 538 AA; 59042 MW; D6B14BE74C3272EF CRC64;

Q24029 Length: 538 May 30, 2002 09:32 Type: P Check: 4193 ..

1 MANDLOYLNA LDVAKTQLYH FTAIVIAMG FETDAYDLFC ISMVTKLGR
51 LYYHDGALK PGSLEPNVSA AVNGVAFCGT LAGQLFFGWL GDRMGRKKVY
101 GMTLMIVIC SIASGLSFGH TPKGVMVTL C FFRWLGFGI GGDYPLSATI
151 MSEYANKKTR GAFAIAVFA MGVGILAGM VAIIVSAFK GAEPAPAYEV
201 DAIGSTVPQA DFVWRITLME GAIPAGLTY YWRMKMPETAR YFALVAKNLK
251 QAANDMSKVL QVEIEAPEK VTAISEAKGA NDFGLFTKEF LRRHGLHLIG
301 TASTWFLDI AFYSQNLFOK DIFSAIGWIP PAQTMNALLE VYKIAARAQTL
351 IALCSIVPGY WFTVAFIDKI GRFAIQLMGF FMTVEMFAL AIDYHMTLK
401 DHRIGFVVMY SFTFEAFNFG PNATTEIVPA EIFPARLRST CHGISAAGK
451 AGAMVGAFGF LYAQTDPDT KTDAGYPPGH WCEELVDRPW LCNFLGMLFT
501 FLVPESNGKS LEDLSRENEG EEEYVAEIRA TSGRTVPV

11AA_SEQUENCE 1.0
ID Q04660 PRELIMINARY; PRT; 938 AA.
AC Q04660;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE A_TM021B04.3 PROTEIN.
GN A_TM021B04.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Dante M., Wamsley P., Gibson A.;
RT "The sequence of A. thaliana TM021B04.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Wash-U;
RT "The A. thaliana Genome Sequencing Project.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Waterston R.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF007271; AAB61068.1; -.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR001320; Ion_glu_t_receptor.
DR InterPro: IPR001311; SBP_glu_t_receptor.
DR Pfam; PF01094; ANF_receptor; 1.
DR Pfam; PF00060; lig_chan; 1.
DR SMART; SM00079; PBpe; 1.
SQ SEQUENCE 938 AA; 105169 MW; 7831B9C30E96CE12 CRC64;

```


004660 Length: 938 May 30, 2002 09:32 Type: P Check: 5621 ..

1 MKRENNLVLS LLEFVIVFLM QVGEAQNRTF NVNNGIVNDI GTAYSNNMILL
51 CINMSLSDFY SSHPETQTRL VTTVDSKND VTAAGATR IPISLNHLAL
101 RYVCVINHOE ICTLLIYIYA FSSKALDLIT NKEVKAILGP WTSMAQAFMI
151 EMGOKSOVPI VTSATSPSL ASIRSQYFFR ATYDSSQVH AIKEIILKLG
201 WREVAIVYVD DTEGEGIMPR LTDVLOEINY RPYRTVISP NATDEISVE
251 LLRMMLPTR VEVVHLVELL ASREFAKATE IGLMKQGYW ILTNTITDVL
301 SIMNETEIEF MGVGLGVKTY VPRSKELENF RSRWTKRFP I SDLNVYGLMA
351 YDATTALALA IEEAGTSNLT FVKMDAKRNV SELQGLGVSQ YGPKLQTLIS
401 RVREOGLAGD FOEINGELOP SVEIIVNVNG QGGRTIGFWM KEYGLFKND
451 QKPASKTTF S WQDRLRPII WPGDTSVPK GWEIPTNGKR LQIGVPVNT
501 FOQFVKATRD PITNSTIFSG FSIDYFEAVI QAIPTYDISD FIFQDGDYD
551 ALVYQVYLGK YDAVADTTI SSNSRMVYDF SLPTYPSGVG LVVPKDSVR
601 RSSTIFLML TLALMLISL SEFIIIGLVW VLEHRVNPDE DGPQYQJLST
651 IFWESFSIMV FAPRERVLSE WARVVIIWY FLVLVLQSY TASLASLIT
701 QHLPTVTNI NSLLAKGESV GYQSSFILGR LRDSGSEAS LVSYSPEHC
751 DALLSKGOAE GVSASVIMEV PYVRIFLGQY CNKYKMVQTP FKVDGLGFVF
801 PIGSPLVADI SRALIKVEES NKANQLENAM FKPIDESCPD PLTNPDNPS
851 VSEFQIGFDS FWVLFVLAAL VCTMALLKFV YQFLKENPNQ RNLRLVMEKE
901 NEPDQKSYIK DYTEKQCCSSG QGMPKNGQEG ANAVNNGN

11AA_SEQUENCE 1.0
ID 022055 PRELIMINARY; PRT; 542 AA.
AC 022055;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN PRT1.
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;
OC Vinceae; Catharanthus.
OX NCBI_TaxID=4058;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. PRETTY IN ROSE;
RA Kai M., Masuda Y., Kikuchi Y., Osaki M., Tadano T.;
RT "Isolation and characterization of a cDNA from Catharanthus roseus
which is highly homologous with phosphate transporter.";
RL Soil Sci. Plant Nutrition 43:227-235(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AB004809; BAA20522.1; -;
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 542 AA; 59580 MW; B7B761A255C8BEDA CRC64;

022055 Length: 542 May 30, 2002 09:32 Type: P Check: 2442 ..

1 MAKEQLOVLN ALDVAKTQWY HFTAIVVAGM GFTDAYDLF CISLVTKLLG

51 RIYVHDGAV KPGTLPPNVS AAVNGVAFEG TLAGQLEFGW LGDKMGRKRV
101 YGMTLMIMVI ASIASGLSFG DKPKAVMATL CFFRFWLFGG IGGDYPLSAT
151 IMSEYANKKT RGAFIAAVEFA MGGFGLAGG MVAITVSASF KAGFPAPAYQ
201 DGAIVASTVPE ADYVWRIITM FGAIPALITY YWRMKMPETA RYTALVAKNA
251 KOANDMSKV LOVELEAEQE KVEKFAQEPN NTFGLFTKEF LKRHGLHLIG
301 TATTWFLDDI AFYSQNLFOK DIFSAIGWIP PAQTMNAIEE VFRIARAQTL
351 IALCSTVPGY WFTVFLIDRI GRFIIQMMGF FMTVFEMFAL AIPYNHWHK.
401 DNRIGFVIMY SLTEFFANFG PNATTFVVA EIPPARLRST CHGISAAAGK
451 AGAIIGAFGE LYAAQPSDPS KTDKGYPPGI GYKNALIVLG CVNFLGVFT
501 FLVPEAKGKS LEEVSKENEE EVENGTELRO QSGHDTRIYV VL

11AA_SEQUENCE 1.0
ID 09ST22 PRELIMINARY; PRT; 536 AA.
AC 09ST22;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. XANTHR-NC;
RA Kai M., Adachi H.;
RT "phosphate transporter.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AB020061; BAA86070.1; -;
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 536 AA; 58909 MW; C2AA6EB8E040504F CRC64;

09ST22 Length: 536 May 30, 2002 09:32 Type: P Check: 2750 ..

1 MAKDLQVLNA LDVAKTQLYH FTAIVVAGM FFTDAYDLFC ISLVTKLLGR
51 IYYHRDGLK PGSLPPNVSA AVNGVAFCGT LAGQLFFGWL GDKMGRKKVY
101 GMTLMMVIC SIASGLSFGH TPKGVMTTLC FFRFWLFGFI GGDYPLSATI
151 MSEYANKKTR GAFIAAVEAM QGFGILAGM VAIIVSAEK GAFPAPTYHA
201 DALASTVSOA DEFWRIITMF GAIPAMTYV WRMKMPETAR YTALVAKNLK
251 QAANDMSKVL QVEIEEEOEK VENVSQNTGN EFGFSEKEL RRHGLHLGT
301 ASTWFLDLIA FYSQNLFOKD IFSAGWIPP AETMNALEEV YRIARAQTLI
351 ALCSTVPGYW FTVEFIDRIG RFAIQLMGEF FMTVFEMFALA IPYHHTWKD
401 NRIGFVIMYS LTFEFANFGP NATTFVVPAP IEPARLRSTC HGISAAGKA
451 GAMIGAFGFL YAAOPTDPK VDAGYPAGIG VRNSLIVLGC VNFLGVETFP
501 LVPESKGKSL EEMSRENEGE EESGTEMKNS GRTVPV

11AA_SEQUENCE 1.0
ID 09FHA7 PRELIMINARY; PRT; 1112 AA.

AC 09FH47;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE GB|AAD55616.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=20181125; Pubmed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
DR EMBL; AB022212; BAB08882.1; -
DR InterPro; IPR000822; Znf-C2H2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
SQ SEQUENCE 1112 AA; 127094 MW; 7C11A579BB26DE85 CRC64;

09FH47 Length: 1112 May 30, 2002 09:32 Type: P Check: 6231 ..

1 MASSFKRDVM AERLYKDAED LIAKGNHIKA LEILEDLISI HMNANDAWLL
51 HIEQGNVEVD LAEKEDRDM EVAYLLGCV A CFSQHTVFE LCAQSLFMLG
101 RNLASVSHYK KCLKAREAL SDSNSKKLSS STPYAFRAIV SRKKDLEGF
151 IEDAESMIAA SKTSPLQKPE PKVRDPEKTP EPSKDDFQRF RSYWVGLDVK
201 IKRDFLKYSI AKLTSEVGGI RKRKAQEALE QALASKKDRK WTEWMCRTKC
251 SKRCSVEEC RKHFEQOHA DFKPSSEKDI VKRIGNNMAR KISLGWEPY
301 DAVAAVEMIK NQADVKAF A SKSNGWSYE WPLAVDEERK QLLQETIKLL
351 VSFCELIKLS GSMRDWMHF PVKYLAKLEV SKQGLIDSHL SETPOSICFL
401 ECHETFRILD FLKTIKCKRN DGTDOYCTAV DSVLGRIQVK VKIDEDPOFS
451 FLDDRRLK INDVQFDDDE GTINVLDP A HYAKAPVHGD DIISWLTDXN
501 SVDKTFPRPI REHNLDIWL A VLKAVQFTCR TLGNKYAKV QVVDYDALT
551 DVENMGVSEN ERKKNLPEDQ WSRVASLLCD VCEERYPKNS LTKLEVRAY
601 RDVFEALHP TLDLFLEDC LNFIREHKS L SDDKVLQ AID LKSVVTQKV
651 LMDTKILLI DNSRISLNN LTRLSAFDNR TYIQLLKP L LNEIIVMES
701 KAKSDAAEAD LNELEKEKL QPKERQSK E KPQSKRRDR TKKRSTSTIS
751 SLDKTVEHK PESTSPSLRT VEEDSMEPED ALASETDMHS RHGEDPLSDH
801 LESAGEAAA RYNSALDMTL KALLSIKIFK EDLMQNSQOI QDQROEQVPT
851 ALQNLFTAVV SEAIQDEGVY SCLSDFLTS QEEFLSMSSD AAKVYVAVLN
901 LMRCKNPER ESLVTRFTL VENKRMSCRK CRRTNSPVQ SSYGIVMAAD
951 SIRELKCAF G NIKFVDILKL IRLGYKMLCD NKKGGCGKKS YVDHIIRCP
1001 PIFTIVLEWE KSETEKEISE TTKALDWEID ISRLYEGL E TNTNYRLVSM
1051 VCGGEGEEH VCLAYEKNRW VNLRRCEFAG EDVGMKNV RFCEGRKVRP
1101 EILYESVRL MA

11AA_SEQUENCE 1.0
ID 09AYT3 PRELIMINARY; PRT; 537 AA.
AC 09AYT3;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN NPTP2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Kai M.;
RT "phosphate transporter.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB042950; BAB21545.1; -
SQ SEQUENCE 537 AA; 59246 MW; DE3D0D68B6C22361 CRC64;

09AYT3 Length: 537 May 30, 2002 09:32 Type: P Check: 5587 ..

1 MAKDQLQVLN ALDVAKTQLY HFTAIYIAGM GFTDAYDLF CISLVTKLLG
51 RIYHHHDGAP KPGLTPNVS AAVNGVAFCG TLAGQLFEFG LGDKMGRKRV
101 YGMTLMMVI CSIASGLSEF HTPKSVWTL CFFRFWLGFG IGDYPLSAT
151 IMSEYANKRT RGAFIAAFA MQFGILAGG MVAIIYSAF KGAPQATYQ
201 TDPLGSTISQ ADFWRIILM FGAIPAMTY YWRMKPETA RYTAIVAKNL
251 KOANDMSKV LQVDIEEQQE KVENVSQNTN NEFGLESKF LRRHGLHLG
301 TASTWFLDI AFYSQNLFOK DIFSAIGWIP PAQTMNALLE YKTIARAQTL
351 IALCSTVPGY WTVTFIDKI GRFAQLMGF FFMVEMFAL AIPYHHWTIK
401 DNRIGFVIMY SLTFFANFG PNATFEVPA EIFPARLST CHGISAAAG
451 AGAMIGARGF LYAAQPTDRK KADAGYPAGI GVRNSLIVLG CVNFLGMVET
501 FLVPESKGS LEEMSRENEG EEESGTEKN SGRTPVPY
11AA_SEQUENCE 1.0
ID 09AYT2 PRELIMINARY; PRT; 523 AA.
AC 09AYT2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN NPTP3.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Kai M.;
RT "phosphate transporter.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB042951; BAB21562.1; -
SQ SEQUENCE 523 AA; 57206 MW; EDC878CD83CFF91A CRC64;

09AYT2 Length: 523 May 30, 2002 09:32 Type: P Check: 7879 ..

1 MSADNNLOVL NALDVAKTQL YHFTTIVYAG MGFTDAYDL FSISIVTKIL
51 GRLYYTKSDL LKPGTLPTV SASVGVVALV GTLAGQLFFG WLGDKMGKRT
101 VYGMTLIMV LCSVASGISF GSTPKGYMAT LCFRFWLGFG GIGDYPPLSA

151 TIMSEYANKK TRGAFIAAVF AMOGFGLFS GIALIVSAG FDHAYNVPSF
201 QENNALSTVP QADYIWRIL MEGALPAFLT YWRMKMPET ARYTALVAKD
251 AKRAQDMGR VLQVEIDPED AKIEQMSRDE TNKGLFSWE FVCRHGLHLF
301 GTCSTWFLLD IAFYSQNLQF KDVSAGVMI PAKTMNAVQ EYKIKARAQT
351 LIALLSTVPG YWFTVAFIDI VGRFAIQLMG FEFMTVEMFA IAIPEHMTK
401 KDNHIGFVAM YALTEFFANF GPNATTFVVP AEIFPARLRS TCHGISAAAG
451 KAGATVGAYG FLYAAQSKDP TKTDAGYPAG IGIKNSLLAL GAINALGMVC
501 TFCVPESKKG SLEEASQETI TEE
11AA_SEQUENCE 1.0
ID 09AYTT1 PRELIMINARY; PRT; 523 AA.
AC 09AYTT1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN NTP14.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RT kai M.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB042956; BAB21563.1; -
SQ SEQUENCE 523 AA; 57242 MW; 9F4C9B6E73CF7DE2 CRC64;
09AYTT1 Length: 523 May 30, 2002 09:32 Type: P Check: 9177 ..
1 MSADNNLQVL NALDVAKTQL YHFTTIVAG MGEFTDAYDL FSISLVTKLL
51 GRLYYTKPEL LKPGTLPTV SASVTGVALV GTLAGQLFEG WLGDKMKRKT
101 YVGMTLIIMV LCSVASGLSF GSTPKGVMAF LCFRFRWLGF GIGDYPISA
151 TIMSEYANKK TRGAFIAAVF AMOGFGLFS GIALIVSAG FDHAYNVPSF
201 QENPALSTVP QADYIWRIL MEGALPAFLT YWRMKMPET ARYTALVAKD
251 AKRAQDMGR VLQVEIDPED AKIEQMSRDK TNKGLFSWE FVRRHGLHLF
301 GTTSTWFLLD IAFYSQNLQF KDVSAGVMI PPAKTMNAVQ EYKIKARAQT
351 LIALLSTVPG YWFTVAFIDI VGRFAIQLMG FEFMTVEMFA IAIPEHMTK
401 KDNHIGFVAM YALTEFFANF GPNATTFVVP AEIFPARLRS TCHGISAAAG
451 KAGATVGAYG FLYAAQSKDP TKTDAGYPAG IGIKNSLLAL GAINALGMVC
501 TFCVPESKKG SLEEASQETI TEE
11AA_SEQUENCE 1.0
ID 065383 PRELIMINARY; PRT; 783 AA.
AC 065383;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F12F1.17 PROTEIN.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Vysotskaya V.S., Osborne B.I., Schwartz J.R., Tortum M., Yu G.,
RA Kwan A., Oji O., Liu S., Buehler E., Conway A.B., Conway A.R.,
RA Dewar K., Feng J., Kim C., Kurtz D., Palm C.J., Li Y., Shinn P.,
RA Sun H., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F12F1 sequence, complete
sequence."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC002131; AAC17615.1; -
DR InterPro: IPR003864; DUF221.
DR Pfam: PF02714; DUF221; 1.
DR SEQUENCE 783 AA; 89528 MW; 2AE3FA091DE5BF5A CRC64;
065383 Length: 783 May 30, 2002 09:32 Type: P Check: 9544 ..
1 MATLIGDIGVA AAINILTAII FLAFAILRI QPFNDRVYFP KWLKIGRSS
51 PLHSGALVSK FVNVNLGSLY RLNMMPAL KMPELIDH AGLDSAVYLR
101 IYLGLKIFV PIALLAWSIL VPVNWTSGL QAKLRNWT SIDIKLSISN
151 IENGSDSLYF GREWTHLVMA YAFTEWTCYV LMKEYEKVAA MRLAFLONEQ
201 RRPDQFTNLG LSOLLQVLY RNVPADPDES ISDSVEHEFL VNHDPHYLTH
251 QVYNNANDLA ALVEQKKSTQ NWLDYQTKY TRNQEHKPRI KTFGLGMGK
301 KVDAIDHYIA EIEKLENGER KVKKDDTSV MPAFVSFKT RMGAASVAGT
351 QOSSDPTWML TEWAPEAREV FWSNLAIPYV SLTVRRLIMH IAFELLTPEF
401 MPIAFVQSL ASIEGIEKNA PFLKSIEND LEKSVIQGFL PGIVLKLFLI
451 FLPSILMWS KREGFVSLSS LERRAARFY IFNLINVELG SVITGAREQ
501 LDSFLKQSAK EIPKTGVAI PIKATFFITY IMVDGWAGIA GEILRLKPLI
551 FFHINKSLV KTEKDREAM NGQINYHAT EPRIQLYELL GLVYAPVYV
601 LLPFIILFFA LAYLVRHQI INVYNQEYES AARFVPDVHG RIISALITAQ
651 ILLMGLLSTK GAAQSTPFL FLPIITFFEH RYCKGRYEP FLRHPLKEAM
701 VKDTLERARE PNPNLKPYLO KAYIHPVKD NDYEDSRFDE ISGYCIEDSD
751 EECVTVPTRK QSRINTPAVS HASRGSRSRP PSK
11AA_SEQUENCE 1.0
ID 022148 PRELIMINARY; PRT; 1374 AA.
AC 022148;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE NON-LTR RETROELEMENT REVERSE TRANSCRIPTASE.
GN AT2G45230.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,

RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RN Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC002387; AAB82639.1; -
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR000477; RVTse.
DR InterPro: IPR000130; Zn_MTpeptidse.
DR Pfam: PF00075; rnaseh; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1374 AA; 158442 MW; 718BA300AE61F581 CRC64;

022148 Length: 1374 May 30, 2002 09:32 Type: P Check: 8740 ..

1 MRILSNWCG VGNTPTVRHL REIRGLYPE VIFLCETKKR RNYLENVVGH
51 LGFEDLHVE PIGKSGGLAL MWKDSVOIKV LQSDKRLIDA LLIQDKEFY
101 LTCIYGPVQ AERGLMERL TRIGLSRSGP WMLTGFENEL VDPSEKIGGP
151 ARKESCLLEF RQMLNSGLW EVNHSYGQFS WYGNRNDLV QCRDRTVAN
201 QAWMELEPQA KATYLOKICS DHPDLINLV GDMWRKWAGF KYDKRWVQRE
251 GFKDLICNF SQQSTKTAL NMEKIASCR EISKWKRVSK PSSAVRIQEL
301 QFKLDAATKQ IPEDRRELAR LKELSQEYN NEQFWQEK RIMMMRNGDR
351 NTKYFHATK NRRANRIQK LIDEGREW SDEDLGRAE AYFKKLFAE
401 DVGTYVEELE NLTPVSDQM NNNLAPITK EEVQATFSI NPHKCPGPDG
451 MNGFLYQOFW ETMGQITEM VQAFERSGSI EEGMNKTNIC LIPKILKAEK
501 MTFDRPISLC NVIYKVIKGL MANRLKKILP SLISETQAAF VKGRLLISDNI
551 LIAHELHLAL SSNNKCSEEF IAIKTDISKA YDRVEMPFL KAMRGLGFAD
601 HMIRLIMECV KSVRYOVLIN GTPHGEIIPS RGLRQGDPLS PLYFVICTEM
651 LVKMLDSAEQ KNQITGLKVA RGAPRISHLL FADDSMFYCK VANDEALGOIT
701 RIIEEYSLAS GQRYVYLKSS IYFGKHISEE RRLVKKRKLK IEREGEQVY
751 LGLPESFOGS KVALTSLYLD RLGKKVLGWQ SNFLSPGCKE ILKAVAMAL
801 PYYMNSCFKI PKTICQOIES VMAEFWMKKN KEGRGLHWKA WCHLSRPKAV
851 GGLGFEKEIEA FNIALIGKQL WRMITKDSL MAKVFKSRYP SKSDPLNAPL
901 GSRPSFAWKS IYEAOVLIKQ GIRAVIGNGE TINVWTDPMI GAKPAKAAQA
951 VKRSHLVSY AANSIHVKD LLDPGRDWN WNLVSLFPD NTQENILALR
1001 PGKETRDRF TWESRSRGY SVKSGYWMT EIIQNRNPO EYLQPSLDPI
1051 FQOIWKLDVP PRIHFLMRC VNNCLSVASN LAYRHLAREK SCVRCPSHGE
1101 TVNHLLEKCP FARLTWAISS LPAPPGGEWA ESLFRNMHV LSVHKSQPEE
1151 SDHALLPWI LWRLLKNRND LVFKGREFTA PQYILKATED MDAMNNRKEP
1201 QPOVTSSTRD RCYKQPPSH GWVGCNTDGA WSKDLGNGCV GWVLRNHTGR

1251 LMLGLRALP SQQSVIETEV EALRWAVLSL SRPNRRVIF ESDSOYLVS
1301 IQNEMDIPSL APRIDIRNL LRHPEEVKFQ FTRREGNNVA DRTARESLSL
1351 MNYPKMYSI TPDWIKNLVD LETV

11AA_SEQUENCE 1.0
ID 082193 PRELIMINARY; PRT; 550 AA.
AC 082193;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE AT2G19880 PROTEIN.
GN AT2G19880.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC005169; AAC62128.1; -
SQ SEQUENCE 550 AA; 62247 MW; 951AA77DAB0EAEL1 CRC64;

082193 Length: 550 May 30, 2002 09:32 Type: P Check: 6852 ..

1 MSTLDSIDAI LFLSLRAFTS PFAVFVQIQ CTICLLALG WLLAEYVRNR
51 EVKRINSIK AGNSLAFLYQ DINELEHSRQ VKLPRVSVM PLKGFGEHNL
101 HNRSQITSL YGPLEFLFV VESTEDPAYH AVSRLSMYQ DHVEAKVVVA
151 GLSTCSQKI HNOLIGVEKM HKDTRYVLEL DDDVRLHPGT IGALTTEMEK
201 NPEVPCSMGF ATGGRTEFLW GGCMMMHADD FRQDRYGVVS GLRDGYSDD
251 MTLASLAGAH KRLITSPVA VEPHPLASDL SFGRYWNYLR KQTFVLESYI
301 SKVNWIMNKA LFAVHCYLSW GFVADYVMAI IHITSALRIY IKGYHOLEDT
351 TSASGGLSVF LMLHIDRRMR FVYGMMLVIT LAICTFIELL SMNNLTREYV
401 QLCNMISPEA PRLSLATYMW GLPFIIGCHI SPMYRGFNKP RTGVFWNVED
451 YPIPNLYPHT IFQNIKSALE NKHFNVEYTP LQQTWSSSIW LSTSLFDRGI
501 LIKPSFTSSA ASWPETMYK GLNGKTSLYL FARAIONIHI SNIAWSARWC
11AA_SEQUENCE 1.0
ID 082362 PRELIMINARY; PRT; 787 AA.
AC 082362;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE AT2G46060 PROTEIN.

GN AT2G46060.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana."
 RT thaliana."
 RN Nature 402:761-768(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005397; AAC62897.1; -;
 DR InterPro: IPR000561; EGF-like.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 SQ SEQUENCE 787 AA; 87537 MW; B6E497738D1D542B CRC64;

082362 Length: 787 May 30, 2002 09:32 Type: P Check: 3197 ..

1 MAEILFLLR TLVFSKALS ICNCLQDAAD NNGNFTVSS FRYPESEVRP
 51 YDTRYIRVDL PPMFSSLNVA IESVDVITAK SISKISKSL PVICFRDGP
 101 PLPDASTNAL QGLELGRFEN GSFERAQDSE IAQCCYPMOK NITRLTNEQ
 151 CLFVLILISY NLDTAQIVRS SAFSESANIS VEGCKTATMW GPFCNQTIYP
 201 LSCSRFDNQT ASVISCADSF PSSCLTGAET KTYALDV DGI AEQLVIMASN
 251 VKVDSNESYL MGYARFEAFA SETLHDYAAD IHKVPILYMK PKGRWYIYI
 301 SLSGRENREA QGTNSSSRVC FSINVKVLGC PVGKAGPNCG QOIYILQAVM
 351 RRGWLTPFQS YFFPVNDASL SGSSTNFPLE PIVSNFSSIP ELDSTWTWYF
 401 LMNIPQGGSG GHIHRLSD STIQEYVYLR FGLPTIDDR DYYYVNRISA
 451 SRSMEFSLYN SSKEMWDFYI LYAREGTWSF GLRQLIDSNT PAASRGSPYL
 501 VSLSLERCPR GCSSYGQCRY AFDANGLTSTY RFFIYLESK ICKHDNGCS
 551 CDRTHGFGDC SIEIVSHQEH IVQSIALIAS NAAALLPAYW ALRQREYEW
 601 VLFSSGSISS ALYHACDVG F WCVLSYNYLQ FMDFWLSFMA VVGTFVYLSL
 651 AGEAVKRTIH TVVAILTALL ALTQATRASN IIVLAIGSL GLLIGFIVEF
 701 VTKYRSYCGS AGFSLNMLDR PRAVKEWESN LIKTLKKRRR WGFVAAGIVA
 751 FTMAAISFKI AIVNHENLAH NGADNYELTR QDSLISRN
 11AA_SEQUENCE 1.0
 ID Q9ZQ76 PRELIMINARY; PRT; 131 AA.
 AC Q9ZQ76;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE AT2G03400 PROTEIN.
 GN AT2G03400.
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana."
 RT thaliana."
 RN Nature 402:761-768(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC006284; AAD17436.1; -;
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 SQ SEQUENCE 131 AA; 14837 MW; 095DCE6344F868D3 CRC64;

09ZQ76 Length: 131 May 30, 2002 09:32 Type: P Check: 6617 ...

1 MVQSQSLSTL TICGSVKVSS LLRNRLNSVK ASSLIGDRCV SCQFLRKSPS
 51 FRSHWKSLLQ RNLLRVEARW PFQGGEGQL DPSSERSESA NEDILIFFEQ
 101 LDLATRVQVK TKCRVFISFK MVANGFDVLTQ Y
 11AA_SEQUENCE 1.0
 ID Q9FRL1 PRELIMINARY; PRT; 654 AA.
 AC Q9FRL1;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE HYPOTHETICAL 73.0 KDA PROTEIN.
 GN F22H5.13.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
 RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RA "Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence."
 RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AC025814; AAG12693.1; -;
 DR HSSP: P16435; 1B1C.
 DR InterPro: IPR001094; Flavdxn-like.
 DR InterPro: IPR001226; Flavodoxin.
 DR Pfam: PF00258; flavodoxin; 1.
 DR PRINTS: PR00369; FLAVODOXIN.
 KW Hypothetical protein.
 SQ SEQUENCE 654 AA; 72997 MW; BADD77B16845516D CRC64;

Q9FRL1 Length: 654 May 30, 2002 09:32 Type: P Check: 7509 ..

1 MSTTSSVRVR LAFVALLSAT TFYCIHKYRR LKHLKNLSLN PSSTLKASRG
 51 KIEFISQGT AKALAQRLE LCASNDIAFD IVDPHSYERE DLPKETLVLF
 101 IASTWDGKP PKNGEFLVNW LGESAEDFRV GSLLSDCKF AVFGVGRAY
 151 GESYNAVAKE LSSRMIGLG LEMIPVGEQ VDDGEIDRAF QDWCDGIVRV

```

201 LKGSAGQETN GVSQIGAVE DDLEYDSTD EDEDNDADG GIVDLEDIAG
251 KAPSKRNGV KATKVDGKKE MTPVIRASL TKOHDFYLLQ GYKIIGSHSG
301 VKICRWTKSQ LRGRCGCYKH SFYGIESHRC METTSLACA NKCVCFCWRHH
351 TNPVGSWMQ KMDEPSVIYK GALDLHKMMI KOMKGPVGT PEKLQEGLN
401 RHCALSLVGE PIMYPEINAL VDELHGRRIIS TFLVTNAQFP EKILMMKPIT
451 QLYVSVDAAV KESLKAIDRP LEADFERFI DSLKALQEKQ QRTVRLTLV
501 KGMNTEELDA YFNLFISIGKP DFIEIKGVTY CGSSATSKLT MENVPWHTDV
551 KAFSEALSLK SNGEYEVACE HAHSCCVLLG RTEKEKVDGK WFTWIDYEK
601 HDLVASGEPE TSTDYMAQTP SWAVYGAQEG GFDPGLRYK KERHPPKPO
651 AVLA

11AA_SEQUENCE 1.0
ID 022301 PRELIMINARY; PRT; 537 AA.
AC 022301;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN MTP1.
OS Medicago truncatula (Barrel medic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_taxid=3880;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98086876; PubMed-9425684;
RA Liu H., Trieu A.T., Blaylock L.A., Harrison M.J.;
RT "Cloning and characterization of two phosphate transporters from
RT Medicago truncatula roots: regulation in response to phosphate and to
RT colonization by arbuscular mycorrhizal (AM) fungi.";
RT Mol. Plant Microbe Interact. 11:14-22(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AF000354; AAB81346.1; -
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 537 AA; 58738 MW; B3BD94029D8C8BC0 CRC64;

022301 Length: 537 May 30, 2002 09:32 Type: P Check: 33 ..
1 MSGELGLVNA LDVAKTQLYH FTTIVIAGMG FFTDAYDLFC ISLVTKLLGR
51 IYTEPNPTR PGTLPSPAQS AVTGVALVGT LAGQLFFGWL GDKLGRKKVY
101 GLTLILMVC SVASGLSPGS SPKSYMATLC FFRFWLGFGI GGDYPLSATI
151 MSEYANKKTR GAFIAAVFAM QGFGILGGCI VALIVASIFD HKYKVPTEEE
201 NPATSLVPO FDIYWRLLIM FGALPALTY YWRMKMPETA RYALVAKNA
251 KQAAADMSKV LOVELEVEEE KVQKMTSDKR NSYGLFSKQF AARHGLALFG
301 TCSTWFLLDI AFYSQNLFOK DIFSAIGWIP PAKEMNAIHE YKIKARAQTL
351 IALCSTVPGY WFTVAFIDHM GRFAIQMMGF FEMTVFMEGL AIPTYDWSKE
401 ENRIGFVVMY SLTEFFSNFG PNAATFVPA EIFPARLRST CHGISAAAGK
451 AGAIVGAFGF LYAAQSKDPT KTDKGYPTGI GIKNSLIMLG VINFGMLCT
501 LLVPEKSKGS LEELSGENEG EGAEATEQEG PREENVA

```

```

11AA_SEQUENCE 1.0
ID 022302 PRELIMINARY; PRT; 533 AA.
AC 022302;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN MTP2.
OS Medicago truncatula (Barrel medic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_taxid=3880;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98086876; PubMed-9425684;
RA Liu H., Trieu A.T., Blaylock L.A., Harrison M.J.;
RT "Cloning and characterization of two phosphate transporters from
RT Medicago truncatula roots: regulation in response to phosphate and to
RT colonization by arbuscular mycorrhizal (AM) fungi.";
RT Mol. Plant Microbe Interact. 11:14-22(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AF000355; AAB81347.1; -
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 533 AA; 58250 MW; 3C4AB00747E4135C CRC64;

```

022302 Length: 533 May 30, 2002 09:32 Type: P Check: 2173 ..

```

1 MSGELGLVNA LDVAKTQLYH FTTIVIAGMG FFTDAYDLFC ISLVTKLLGR
51 IYTEPNPTR PGTLPSPAQS AVTGVALVGT LAGQLFFGWL GDKLGRKKVY
101 GLTLILMVC SVASGLSPGS SPKSYMATLC FFRFWLGFGI GGDYPLSATI
151 MSEYANKKTR GAFIAAVFAM QGFGILGGCI VALIVASIFD HKYKVPTEEE
201 NPASLVLPO FDIYWRLLIM FGALPALTY YWRMKMPETA RYALVAKNA
251 KQAAADMSKV LOVELEVEEE KVEKMTSDKR NSYGLFSKQF AARHGLALFG
301 TCSTWFLLDI AFYSQNLFOK DIFSAIGWIP PAKEMNAIHE YKIKARAQTL
351 IALCSTVPGY WFTVAFIDHM GRFAIQMMGF FEMTVFMEAL AIPTYDWSKE
401 ENRIGFVVIY SLTEFFANFG PNAATFVPA EIFPARLRST CHGISAAAGK
451 AGAIVGAFGF LYAAQSKDPT KTDKGYPTGI GIKNSLIMLG VINFGMLCT
501 LLVPEKSKGS LEELSGENEG EGAEATEQEG SRV

11AA_SEQUENCE 1.0
ID 022548 PRELIMINARY; PRT; 538 AA.
AC 022548;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER.
GN LEPT1.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu C., Muchhal U.S., Mukatira U., Kononowicz A.K., Raghothama K.G.;
RT "Tomato phosphate transporter genes are differentially regulated in
RT plant tissues by phosphorus.";
RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.

```

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: AF022873; AAB82146.1; -
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SO SEQUENCE 538 AA; 58699 MW; 9B1961E0CB7EC0A6 CRC64;

022548 length: 538 May 30, 2002 09:32 Type: P Check: 4793 ..

1 MANDLOVLNA LDVAKTQLYH FTAIVAGMG FFTDAYDLFC ISMTKLIGR
 51 LYHHHDGALK PGSLPPNWSA AVNGVAFCGT LAGQLFFGWL GDKMGRKKVY
 101 GMTLMIMVIC SIASGLSEGH TPKGVMTLIC FFRFWLGGI GGDYPLSATI
 151 MSEYANKTR GAFAAVFAM QGFGILAGM VAIIVSAFK GAFFAPAYEV
 201 DAIGSTVPOA DEVRILIMF GAIPAGLTY WRMKMPETAR YTALVAKNLK
 251 QAANDMSKVL QVEIEAPEK VTAISEAKGA NDFGLFTKEF LRRHGLHLG
 301 TASTWFLDI AFYSQNLFOK DIFSAIGWIP PAQTMALKEE YKIKARQTL
 351 IALCSTVPGY WFTVAFIDKI GRFAIQLMGF FFMVTMEMAL AIPYHMTLK
 401 DHRIGEVVMY SFTFFANEG PNATTFVVA EIFPARLRST CHGISAAGK
 451 AGAMVGAFCF LYAAQPTDPT KTDAGYPPGI GVRNSLIYLG CVNPLGMLFT
 501 FLVPESNGKS LEDLSRENEG EETVAEIRA TSGRTVPV

11AA_SEQUENCE 1.0

ID 022549 PRELIMINARY; PRT; 528 AA.

AC 022549;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE INORGANIC PHOSPHATE TRANSPORTER.
 GN LEPT2.

OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_Taxid=4081;

RN [1]
 RP SEQUENCE FROM N.A.

RA Liu C., Muchhal U.S., Mukatira U., Kononowicz A.K., Raghothama K.G.;
 RT "Tomato phosphate transporter genes are differentially regulated in
 plant tissues by phosphorus";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: AF022874; AAB82147.1; -
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.

KW Transmembrane.
 SO SEQUENCE 528 AA; 57762 MW; 3B165AE6E2DE488C CRC64;

022549 length: 528 May 30, 2002 09:32 Type: P Check: 3242 ..

1 MAVGDNDNNN LQVLNLDLIA KTQLYHTAI VIAGMGFTD AYDLPSISLV
 51 TKLLGRLYYT KPDLLKPGTL PPAVSASVTG VALVGTLAGQ LFFGWLGDKM
 101 GRKKVYGWTL VLMVCSVAS GLSFGSTPKG VMTTLCHFRF WLGFGIGGDY
 151 PLSATIMSEY ANKKTGCAFI AAVFAMQFG ILFSGIYALI TAAGFDHAYR
 201 SPTFEENAL STVPOSYIW RIIMFGSLP AALTYWRMK MPETARYTAL
 251 VAKDAKRAAQ DMGKLOVEI ESEAKIEQI SRNETNQFGL FSWEFVRRHG

301 LHLFGTGSTW FLIDIAFYSQ NLFQKDVESA VGWIPKAPTM NAVQEVYKIA
 351 RAQTLIALCS TVPGYWFTVA FIDIIGRAFI QLMGFFEMTV FMAFAIPYH
 401 HWTLEANRIG FIVMYSLTFF FANFGPNAT FVPAEIRPA RLSTCHGIS
 451 AAAGKAGAIV GAYGFLYAAQ SKDPNKTDAQ YPAGIGIKNS LIVLGICNAL
 501 GMLCTFCVPE PKGKSLSEAS QETITGEA

11AA_SEQUENCE 1.0

ID 09XEL6 PRELIMINARY; PRT; 396 AA.

AC 09XEL6;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HIGH-AFFINITY PHOSPHATE TRANSPORTER PT1 (FRAGMENT).
 OS Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Poaceae; Triticum.
 OX NCBI_Taxid=4565;

RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=CV. CHINESE SPRING;
 RA Li Y.J., Chen S.Y., Li Z.S.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

DR EMBL: AF110180; AAD26146.1; -
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.

FT NON_TER 1 1
 FT NON_TER 396 396
 SO SEQUENCE 396 AA; 43519 MW; A2ACD6A657E0B25E CRC64;

09XEL6 length: 396 May 30, 2002 09:32 Type: P Check: 7287 ..

1 FFTDAYDLFC ISLVTKLIGR IYYREGADA PGSLPPNLAA AVNGVAFCGT
 51 LSGQLFFGWL GDRMGRKKRY GMTLMCMVLC SIASGLSEGS TPGSVNATLC
 101 FFRFWLGGI GGDYPLSATI MSEYANKTR GGFIAAVFAM QGFGILTGV
 151 VTLIVSAFR AAFPTQAYOD APLASTPAQA DEVRILIMF GAVPALMTYY
 201 WRMKMPETAR YTALVAKNAK QAAADMSKVL QVDIGAEED PKANDVGAA
 251 DDRNSFGLFS GEFLRRHGL LIGTATCWF LIDIAFYSQNL FOKDIFTAIN
 301 WIPKAKTMSA LEEVHRIARA QTLIALCGTV PGYWFTVALI DRIGFEWQL
 351 GGFEMAVER LGLAFPYHHW TTPGNHIGFV VLYALTFEFA NFGPNA

11AA_SEQUENCE 1.0

ID 09ZPF4 PRELIMINARY; PRT; 382 AA.

AC 09ZPF4;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE T3H13.9 PROTEIN.

GN T3H13.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_Taxid=3702;

RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA WashU;
 RT "The A. thaliana Genome Sequencing Project.";


```

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RT Drone K., Nguyen C.;
RL "The sequence of A. thaliana T3H13."
RN Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Waterston R.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AF128396; AAD17374.1; -
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 382 AA; 41929 MW; 979C03576443F8D7 CRC64;

Q9ZPF4 Length: 382 May 30, 2002 09:32 Type: P Check: 2196 ..

1 MVRFHGCSR RYGFETDSYD LEVLSITKL LGRIVYQVPG SSSPSGLPDG
51 ISVAVSGVAF AGTFLGQIF GCLGDKLGRK RYGLTLILM TICSIASSLS
101 FGKDKTVMV TLCEFRFWLG FGIGDYPLS ATIMEYANK RTGAFIASV
151 FAMGVGIIA AGVSLLSVY LFEIEFPSRA YILDGAATV PQADYVWRII
201 LMVGLPALI TYWRMKMPE TARYTALVAK NAEQAALDMN KEITFEVDEK
251 REALTICQSE NCLIDRLGRY WIOIIGFMM TVFMLVLAVP YHHWTLPLANR
301 IGFVIFYSFT FFFSNFGPNA TTFIVPAEIF PARIRSTCHG ISAASGAGA
351 MVGSFGFAAL GISLEELTGE TKPERVKEKI GV

11AA_SEQUENCE 1.0
ID Q9SWG9 PRELIMINARY; PRT; 540 AA.
AC Q9SWG9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER.
GN PT1.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NAMJAK;
RA Yun S.J., Baek S.H., Lee M.C.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AF156695; AAD38859.1; -
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 540 AA; 59150 MW; 3593F4335A8C8568 CRC64;

Q9SWG9 Length: 540 May 30, 2002 09:32 Type: P Check: 8707 ..

1 MANDLQVLNA LDVAKTQLYH FTAIVIAGMG FETDAYDLFC ISWTKLLGR
51 IYHHDNALK PGSLPPNVA AVNGVAFCGT LAGQLFFGWL GDKMGRKKVY
101 GMTLMIMVIC SIASGLSFGH TPKSVMTTLC FFRFWLGFGI GG DYPLSATI
151 MSEYANKKTR GAFLAAVFAM OGFGILAGGM VAIIVSSAFK GAFFAPAYEV

```

```

201 DALASTVSQA DFWRIILME GAIPAGLTYT WRMKPETAR YTALVAKNLK
251 QAANDNSKVL QVEIEAPEK VGAISEAKGA NEFGLESKEF LRRHGLHLIG
301 TASTWFLLDI AFYSQNLFOK DIFSAIGWIP PAQTMNALLEE YKIKARAQTL
351 IALCSTVPGY WFTVAFIDRI GREAIQLMGF FEMTVEMFAL ALPYHMTLK
401 DNRIGFVVMY SLTFEEANFG PNATTFVVA EIPPARLRST CHGISAAAGK
451 AGAMVGAFCF LYAAQPTDPK KTDAGYPPGI GVRNSLIVLG CVNFLGMLFT
501 FLVPESNGKS LEEMSGRENEG EETVAEMRA TSGRTVLKFK

11AA_SEQUENCE 1.0
ID Q9LIS5 PRELIMINARY; PRT; 537 AA.
AC Q9LIS5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER.
GN PT1.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Yun S.J., Baek S.H., Lee M.C.;
RT "Differential regulation of tobacco phosphate transporter genes by
RT phosphorus."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF156696; AAF74025.1; -
SQ SEQUENCE 537 AA; 59198 MW; 75820D1C35C2CF01 CRC64;

Q9LIS5 Length: 537 May 30, 2002 09:32 Type: P Check: 6167 ..

1 MAKDQLVLN ALDVAKTQLY HFTAVIAGM GFETDAYDLF CISLVTKLIG
51 RIYVHHDGAP KPGLPPNVS AAVNGVAFCG TLAGQLFFGW LGDKMGRKRV
101 YGMTLMIMVY CSIASGLSFG HTPKSVMTTL CFFRFWLGFG IGGDYPLSAT
151 IMSEYANKKT RGAFIAVFA MGFEGILAGG MVAIIVSAF KGAFPAQTYQ
201 TDPLGSTVSQ ADFVWRIILM FGAIPAAPTY YWRMKMPEA RYTALVAKNL
251 KOANDNSKV LQVDIEEQE KVENVSQNTN NEFGLESKEF LRRHGLHLIG
301 TASTWFLLDI AFYSQNLFOK DIFSAIGWIP PAQTMNALLEE YKIKARAQTL
351 IALCSTVPGY WFTVFEIDKI GREAIQLMGF FEMTVEMFAL AIPYHMTLK
401 DNRIGFVIMY SLTFEEANFG PNATTFVVA EIPPARLRST CHGISAAAGK
451 AGAMIGAFCF LYAAQPTDRK KADAGYPAGI GVRNSLIVLG CVNFLGMVFT
501 FLVPESNGKS LEEMSGRENEG EESGTEMKN SGRTPVPY

11AA_SEQUENCE 1.0
ID Q9M5K0 PRELIMINARY; PRT; 265 AA.
AC Q9M5K0;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE PHOSPHATE TRANSPORTER (FRAGMENT).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;

```



```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. INDICA;
RA Yu F., Zhang A., Zhang F., Chen S.;
RT "Rice phosphate transporter.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF229169; AAF40188.1; -.
FT NON_TER 1 1
FT NON_TER 265 265
SQ SEQUENCE 265 AA; 28329 MW; DC88DD1D52E06A74 CRC64;

Q9M5K0 Length: 265 May 30, 2002 09:32 Type: P Check: 8753 ..

1 FFTDAYDLFC ISLVSKLLGR IYTDIASDT PGSLPPNVA AVNGVALCGT
51 LAGQLFFGWL GDKLGRKSVY GFTLVLMVVC SVASGLSFGK TAKGVVATLC
101 FFRFWLGFGL GGDYPLSATI MSEYANKRTR GAFIAVFEAM QGFGLFGAI
151 VALVVSAGFR NAYPAPSYAD GRAASLVPEA DYVWRILMF GTRPGAQLL
201 LAHENARRFA RYTGSAQTP SKAGRXHXG VTXRXFOXGX XAPGXRGVXR
251 RXXPRXXXKG XFXG

11AA_SEQUENCE 1.0
ID Q9M562 PRELIMINARY; PRT; 393 AA.
AC Q9M562;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE PHOSPHATE TRANSPORTER (FRAGMENT).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. JAPONICA;
RA Yu F., Zhang A., Zhang F., Chen S.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF239619; AAF42956.2; -.
FT NON_TER 1 1
FT NON_TER 393 393
SQ SEQUENCE 393 AA; 42860 MW; E45E1E7A5A042B9B CRC64;

Q9M562 Length: 393 May 30, 2002 09:32 Type: P Check: 15 ..

1 FFTDAYDLFC ISLVSKLLGR IYTDIASDT PGSLPPNVA AVNGVALCGT
51 LAGQLFFGWL GDKLGRKSVY GFTLVLMVVC SVASGLSFGK TAKGVVATLC
101 FFRFWLGFGL GGDYPLSATI MSEYANKRTR GAFIAVFEAM QGFGLFGAI
151 VALVVSAGFR NAYPAPSYAD GRAASLVPEA DYVWRILMF GTVPALITYY
201 WRMKMPETAR YTALIRNAK QAAADMSKYL DTEIQEDADR AEAVAGGAG
251 NEWGLFSRHF VRRHGVHLVA TTSTWFLDI AFYSQNLFOK DIFSVCWIP
301 PARTMANVEE VERIARAQAL IALCGTIPGY WFTVAFIDVA GRAFIQLMGF
351 AMMTVFMLGL AAPYHHWTP GNHTGFVVMY GTTFEFANFG PNA

11AA_SEQUENCE 1.0
ID Q9MAX4 PRELIMINARY; PRT; 709 AA.
AC Q9MAX4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE TRANSPOSASE.
OS Zea mays (Maize).
```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=JITTERY;
RA Yan X., Li J., Cowperthwaite M., Fu H., Dooner H.;
RT "Jittery, a low-copy, Mu-related transposon apparently mobilized by
RT BSMV infection.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF247646; AAF66982.1; -.
SQ SEQUENCE 709 AA; 81941 MW; F1080D806D982ABF CRC64;

Q9MAX4 Length: 709 May 30, 2002 09:32 Type: P Check: 7892 ..

1 MISHLFFSWQ SSSSFVPDCE HSLKPVIGMS FDSLDELEGF YKTYAHBCGF
51 SVRIGAQGKK NDVEHKRFV CSREGFTRRC AEAKNQKKHF ETRCGCNARV
101 YVRLGQDKRY YIASFVEEHN HGLVSPDKIP FLRSNRTICQ RAKTTLFTCH
151 KASIGTSQAY RLLQVSDGFD NIGCMKRDLQ NYRGLREKI KNADAQLFVA
201 QMERKKEANS AFFYDFAVDE HGKLVYICWA DATCRKSYTH FGDLSYDAT
251 YSTNQYNMRF APFTGVNHHM QRVFFGAFL ANEKIESYEW LFRTFLVAMG
301 GKAPRLIITD EDASIKSAIR TTLPTIHLR CMWHIMEKVS EKVGHPTSHD
351 KEFWDALNTC VMGSETPEEF EMRWNALMDA YGLESNEWLA NRYKIRSWI
401 PAFEMDTPLA GVLRTTSRSE SANSFNRFI HRKLCFVEFW LRFDTALERQ
451 RHEELKADHI SIHSTPLVRT PWVEKQASI LYTHKVEKIF QEEVIAARDH
501 CSVLGTQOD AVKEVVVSDG SMRDVVQWC TSNIFGRSC KLFEKIGIPC
551 CHIILAMGE KLYELPSSYI LKRWETRCKR ECVYDDGDL LEEKPEDANE
601 AEKRRKITVV RNKIEAIDR AKSSNEAMDF LVSSYLANICE SLGHIVSMV
651 QPTQEEYENF IGCKIPADIQ IHPNDVRSK GRSKRIRAK ELPKSRGKN
701 ARKDMEAPL

11AA_SEQUENCE 1.0
ID Q9LKL5 PRELIMINARY; PRT; 390 AA.
AC Q9LKL5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER (FRAGMENT).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JINGX117, AND CV. JAPONICA; TISSUE=SHOOT;
RA Yu F., Zhang A., Zhang F., Chen S.;
RT "Rice shoot phosphate transporter.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF271893; AAF76345.1; -.
FT NON_TER 1 1
FT NON_TER 390 390
SQ SEQUENCE 390 AA; 43066 MW; 61A281BB288366E1 CRC64;

Q9LKL5 Length: 390 May 30, 2002 09:32 Type: P Check: 9448 ..

1 FFTDAYDLFS ISLVTKLLGR IYFNPASKS PGSLPPNVA AVNGVAFCGT
```

51 LAGQLEFGWL GDKMKRKKVY GMTLMLMVIC CLASGLSFGS SAKGVNATLC
101 FFRFNLGFGI GGDYPLSATI MSEYANKRTR GAFIAAVFAM QGFGNLGTGI
151 VAIIVSAAFK SRFDAPAYRD DRIGSTVPKP TTRGASCSGS APIPALLTTY
201 WRMKMETAR YTALVAKNAK QAAADMTQVL NVEIVEONK ADEVAREQF
251 GLFSRQFLRR HGRHLGTTV CWFVLDAFY SSNLFQKDIY TAVQMLPKAD
301 TMSALEMEK ISRAQTLVAL CGTIPGYWT VEFIDIIGRF VIQGGFFFM
351 TAFMLGLAVP YHHWTPGNH IGFVVMYAFY FFFAXXGPXA

11AA_SEQUENCE 1.0
ID Q9ATE1 PRELIMINARY; PRT; 323 AA.

AC Q9ATE1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE PHOSPHATE TRANSPORTER JX17 (FRAGMENT).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_Taxid=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Ming F., Shen D.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335588; AAK25766.1; -.
DR InterPro; IPR002173; PfkB.
DR PROSITE; PS00584; PFKB_KINASES_2; UNKNOWN_1.
FT NON_TER 1 1
FT NON_TER 323 323
SQ SEQUENCE 323 AA; 34038 MW; 17688B53225530F4 CRC64;

Q9ATE1 Length: 323 May 30, 2002 09:32 Type: P Check: 2612 ..

1 LFCISLVTKL LGRIYYRVDS SPSPTGLRPH VSASVNGVAF VGTLSGQLFF
51 GWLGDKLGRK RYVGITLMLM VLCSLASALS FGHPTPSVMA TLCPFRFWLG
101 FGIGDYPLS ATIMSEYANK KTRGAFIAAV FAMQFGIIT GGLVAIIVSA
151 SFRAFPAPP YGEDPYASTP PQADFWIRII LMLGALPAAL TYYWRTKME
201 TARYTALVAN NAKQAAADMS KVLQVEMRN IGNGSGSRA VRAYLRVCP
251 AARAABGCHV GDVVAAGHCY LQPEVPBGH IQRGVDVPOG GDDERAGAV
301 PHRAGADADR AVRDCARLLV HGG

11AA_SEQUENCE 1.0
ID Q9AXE0 PRELIMINARY; PRT; 333 AA.

AC Q9AXE0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PHOSPHATE TRANSPORTER (FRAGMENT).
OS Oryza rufipogon (Wild rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_Taxid=4529;
RN [1]
RP SEQUENCE FROM N.A.
RA Ming F., Shen D., Zhang Y.;
RT "Common wild rice phosphate transporter mRNA.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF337531; AAK06857.1; -.
FT NON_TER 1 1
FT NON_TER 333 333

SQ SEQUENCE 333 AA; 36461 MW; 9C95C37D45BF0EB0 CRC64;

Q9AXE0 Length: 333 May 30, 2002 09:32 Type: P Check: 5023 ..

1 LFCISLVTKL LGRIYYFNPA SKSPGSLPN VSAAVNGVAF CGTLAQOLF
51 GWLGDKMKRK KYGMITLMLM VICCLASGLS FGSSAKGVMA TLCPFRFWLG
101 FGIGDYPLS ATIMSEYANK RTRGAFIAAV FAMQFGNLT GGIIVIVSA
151 AFKSRFDAPA YRDDRAGSTV PQADYAMRIV LMFCAIPALL TYYWRTKME
201 TARYTALVAK NDKKAAADMA RVLNVELVDE QEKAAATAA AAEEEARRE
251 QYGLFSREFA RRHGHHLGT TVCWFVLDDIA YYSQNLQKD IFSALGIWE
301 AKTMSALDEL YHIRAQTLI ALCGTVPGYW FTV

11AA_SEQUENCE 1.0
ID Q9M1T0 PRELIMINARY; PRT; 535 AA.

AC Q9M1T0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORT PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Obermaier B., Ottenwaelder B., Duchemin D., Zeitler K., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quettler F., Salanoubat M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL138650; CAB77590.1; -.
SQ SEQUENCE 535 AA; 58332 MW; 8FA274A5413FF800 CRC64;

Q9M1T0 Length: 535 May 30, 2002 09:32 Type: P Check: 3263 ..

1 MAGDQLNVLN ALDVAKTQWY HFTAILIAGM GFTDAYDLF CISLVTKLLG
51 RIYYHVDGSE KPGTLPPNVS AAVNGVAFPG TLAGQLFFGW LGDKLGRKKV
101 YGMTLMVAVL CSIASGLSFG SNPKTYWTL CFFRFWLFGG IGDYPLSAT
151 IMSEYANKKT RGAFIAAVEA MQFGILTGG IFALIVSAF EAKFPAPTYQ
201 IDALASTVPQ ADYVWRITLM VGALPAAMTY YSRSKMPETA RYTALVAKDA
251 KLAASNMKV LQVEIEAQO GTEDKSNSFG LFSKEFMKRH GLHLGTST
301 WFLLDIAFYS QNLFOKDIFS AIGWIPPAQT MNAIQEVFKI ARAQTLIALC
351 STVPGYWFIV AFIDVIGRA IQMMGFFMT VFMFALAIPY DHWTNENRI
401 GFVAMYSLTF FFANFGPNAT TFFVPAETFP ARFRSTCHGI SAASGLGAM
451 VGAFGFLYLA QSPDKTKEH GYPGIGYKN SLIVLGTVNL LGMVFTLLVP
501 ESKGKSLEEM SGENEQNDES SSSSNNNSNN AVSTA

11AA_SEQUENCE 1.0
ID Q9LZ39 PRELIMINARY; PRT; 428 AA.

AC Q9LZ39;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 48.9 KDA PROTEIN.
GN T22P11_220.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
KW EMBL; AL162971; CAB86000.1; -
SQ Hypothetical protein.
SEQUENCE 428 AA; 48883 MW; 425E34860C0E1E18 CRC64;

Q9LZ39 Length: 428 May 30, 2002 09:32 Type: P Check: 3770 ..

1 MTLPLFLAAV FVLQLSLTLT VAEIKSFITIS NDSRPVILL E KFGIIEIGHV
51 TVSVSSVSVL SPILDSKLG FVLSEESLP HVLLELQONF SECVLDSHYI
101 LHEFTFVDLS PPRSQFSKS YPITSPNDYS LEFANCVPET RVSMKVHTEI
151 YHDLYPNGSR DYLLAGSAQL PGLYLVFFLC YLSFLCFWLC PCWNHKQIVK
201 RIHLMLTALL LVKSLTLICA AVYKHVYKVT GTAHGWNIVF YIFQFISVVL
251 LEMVIVLIGN GMSFLKPKLH VKEKKLLIV VPLQVLANIA SIVIGETGPY
301 TQDWVSWNOI FFLADITCCC AIVFAMWSM CCLRETSKTD GKAVKNLAKL
351 PVLKRFYVLV IGYLFFTRIV VVVMKMKADF TYQVSNAAE EIALTSFYCL
401 MFYMFRIEK NEYCDVDDEE EIVELSLK
!!AA_SEQUENCE 1.0
ID Q9ASF0 PRELIMINARY; PRT; 794 AA.
AC Q9ASF0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 85.9 KDA PROTEIN.
GN P0456F08.9.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0456F08.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AP002901; BAB39409.1; -
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;
transferase.

SQ SEQUENCE 794 AA; 85935 MW; 289C9C2539F1A1AC CRC64;

Q9ASF0 Length: 794 May 30, 2002 09:32 Type: P Check: 5843 ..

1 MYDAGLAIPP LPTTSPVFP PPIRSYPLP ANKPYPNRPA VSPASIHPT
51 NHGKAHGVP I AAHSKERHHH SMLVNTNTHN THAGPVAPP KGRHHSLPV
101 NNTRVKGPAY SPNSPSIHR KHGIPVAAP KOHSSNLPPS HHRPHKGSFP
151 VISPTPHKAD NASATKHGRS GLHSPAPAP VGLPPSEGNA RGNPAYAPRH
201 PHEXSPSNS PELAQDIASG VLMKQSVRV MGNAAATEDP EKTVVLLIDLV
251 PLGEKFDKAT ALVFERFWH KOVNOFTSSV CVCVCVSLPP SPPTAPRMNN
301 GLSNVNDPRL HPLAVDVGNH RETKSRGIIV IIVLSSVFAF ILCGAAIYI
351 CFKIRNRNHL TEESPMPPKP AGPGSAVVG S RUGSRPISAS PSFSSSIYTY
401 KGTAKTESLI EMERATQRFD NSRIEGEGF GRVYEGILED GERVAVKILK
451 RDDQVTRF LAELEMLSRL HHRNLVKLIG ICTEEHICRL VVELVPNGSV
501 ESHLHGSDBG TAPLDWDARL KIALGAARAL AYLEDSSPR VIHRDFKSSN
551 ILLEHDEFTPK VSDEGLARTA IGEENEHIST RVMGTFGYVA PEYAMTGHL
601 VKSDVYSYGV VLELLTGRK PVDILRPPGQ ENLVAMACPF LTSRDGLETTI
651 IDBSLGNLIL FDSIAKVAAI ASMCVQPEVD QRPFMGEVWQ ALKLVCEGS
701 EFNESGSFSQ DHIQDSGII SRASLDVDE PVVSAELFNA SAHYDTLDAS
751 GSFRRYSSSG PLRVGRTGHN RERGLSTGSS SEHCQTQRF R IDSE
!!AA_SEQUENCE 1.0
ID Q9ARI9 PRELIMINARY; PRT; 543 AA.
AC Q9ARI9;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER 1.
GN PTL.
OS lupinus albus (white lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genistaceae; lupinus.
OX NCBI_TaxID=3870;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu J., Uhde-Stone C., Li A., Vance C.P., Allan D.L.;
RT "A phosphate transporter with enhanced expression in proteoid roots of
RT white lupin (Lupinus albus L.).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY026321; AAK01938.1; -
SQ SEQUENCE 543 AA; 59687 MW; D754FC2C7DA515F4 CRC64;

Q9ARI9 Length: 543 May 30, 2002 09:32 Type: P Check: 1485 ..

1 MAREHIHVLN ALDVAKTQWY HFTAIITAGM GFTDAYDLF CISLVTKLIG
51 RIYHVVDGAG KPCTLPPNVS AAVNGVAFCG TLSQLFFGW LGDKMGRKKV
101 YGMTLIMIVI CSIGSLSGF HSPNSVLATL CFERFWLFGF IGDYPLSAT
151 INSEYSNKT RGSFIAVFA MGGFGLGGG IFALIISAF KARFADAPAVE
201 VDPVGSTVPQ ADYIWRITVM VGALLPALTY YWRMKMPETA RYTAIVAKNA
251 QQAADMSKV LOVEIOSETN KEAOGKPSF GLFSKEFLRR HGLHLGTAG
301 TWFLLDIAFY SONLFQKDIF SAIGWIPPAK TMNALDEVYR IARAQTIAL

```

351 CSTVPGWFT VALIDRGRF AIQLMGFFEM TVFMALAI P YDHWTHKDN R
401 IGFVIVISLT EEFANFGPNA TTFVPAEIT PARFRSTCHG ISSASGLGA
451 IVGAFGLYL AQNKDKSKTD AGYPAGIGV NSLIVGVN ILGCFETFLV
501 PEPNGKSL EE MSGENEEEP TKEGSYNMNN NNNNNNRTV PYV

11AA_SEQUENCE 1.0
ID Q40115 PRELIMINARY; PRT; 214 AA.
AC Q40115;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RIBONUCLEASE (RNase LC1) PRECURSOR.
OS Lufla cylindrica (Smooth loofah) (Sponge gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; lufla.
OX NCBI_TaxID=3670;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T., Sasaki K., Funatsu G.;
RT "Cloning and sequencing of two cDNA clones coding for ribonucleases
RL from the seeds of sponge gourd (lufla cylindrica).";
RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; D64011; BAA10891.1; -.
DR InterPro; IPR001568; RNase_T2.
DR Pfam; PF00445; ribonuclease_T2; 1.
DR PROSITE; PS00531; RNase_T2_2; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 214 RIBONUCLEASE (RNase LC1).
SQ SEQUENCE 214 AA; 24080 MW; 0FE50D18D5D358BE CRC64;

Q40115 Length: 214 May 30, 2002 09:32 Type: P Check: 3380 ..

1 MAMAKREIVL VFVLTILFPM VKSOTFDSFW MVQHWPPAVC SFOQGRVCVQ
51 GURSTIHGV WPQKGTSVI NCPGPTFDT KISHLESTLN VDMPNVITGN
101 NKWFWGHEWN KHGICSVSKF DQDAYFQMAI NMRNSIDLIS ALRVGVVPRN
151 GRSKARQRVQ SAIRAQLGKE PVLRCRGTR QSRLLLEIVMC FDDDGVTLLN
201 CNPANSNCPN SFIF

11AA_SEQUENCE 1.0
ID Q04381 PRELIMINARY; PRT; 521 AA.
AC Q04381;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE PROTON/PHOSPHATE COTRANSPORTER.
GN ATP4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97375319; PubMed=9235602;
RA Lu Y.-P., Zhen R.-G., Rea P.A.;
RT "ATP4: A fourth member of the Arabidopsis phosphate transporter gene
RT family (Accession No. U97546) (PGR97-082).";
RL Plant Physiol. 114:747-747(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; U97546; AAB69122.1; -.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.

```

```

KW Transmembrane.
SQ SEQUENCE 521 AA; 57356 MW; 62DF9254EC6942D5 CRC64;

Q04381 Length: 521 May 30, 2002 09:32 Type: P Check: 4865 ..

1 MADQQLGLK ALDVAKTQLY HPTAIVAGM GFTDAYDLF CVSLVTKLIG
51 RLVEFNPTSA KRGSLPHVA AAVNGVALCG TLGQLFFGW LGDKLGRKKV
101 YGITLIMIL CSVASGLSLG NSAKGVTTL CFFRWLFGG IGDYPLSAT
151 IMSEYANKKT RGAFTAAVFA MGVRLIAGG FVALAVSSIF DKKFSPTYE
201 QDRFLSTPQQ ADYIWRITVM FGALPALTY YWRMKMPETA RYVALVAKNI
251 KQATADMSKV LQTDLELER VEDVKDPKK NYGLFSKEFL RRHGLHLGT
301 TSTWFLDLIA FYSQNLFOKD IFSALGWIPK AATMAIHEV FKARAQTLI
351 ALGSTVPGYW FTVAFIDIIG RPAIQLMGFF MMTVEMFAIA FPYNHWILPD
401 NRIGFVVMYS LTFEFANFGP NATFIVPAE IFPARLRSTC HGISAATGKA
451 GAIVGAFGL YAAQPODKTK TDAGYPPGIG VKNSLIMLGV INFGMLFTF
501 LVPEPKGKSL EELSGAEVD K

11AA_SEQUENCE 1.0
ID Q43650 PRELIMINARY; PRT; 540 AA.
AC Q43650;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER 1.
GN STPT1.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. DESIREE; TISSUE=ROOT;
RX MEDLINE=97246321; PubMed=9090882;
RA Leggewie G., Willmitzer L., Riesmeyer J.W.;
RT "Two cDNA's from potato are able to complement a phosphate uptake
RT deficient yeast mutant: Identification of phosphate transporters from
RT higher plants.";
RL Plant Cell 9:381-392(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; X98890; CAA67395.1; -.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 540 AA; 59108 MW; C79B3E14933763BF CRC64;

Q43650 Length: 540 May 30, 2002 09:32 Type: P Check: 8926 ..

1 MANDLQVLNA LDVAKTQLYH FTAIVIAGMG FFTDAYDLFC ISMTKLLGR
51 IYHHDNALK PGSLLPNVSA AVNGVAFCGT LAGQLFFGWL GDKMGRKKVY
101 GMTLMIMVIC STASGLSFGH TPKSVMTTLC FFRFWLFGI GGDYPLSATI
151 MSEYANKKTR GAFTAAVFA MGVRLIAGG VALIVSSAFK GAFPAPAYEV
201 DALASTVSQA DEYWKRIIMF GALPAGLTYY WRMKMPETAR YTALVAKNLK
251 QANDMSKVL QVEIEAPEK VAAISVANGA NEFGLFSKEF LRRHGLHLIG
301 TASTWFLLDI AFYSQNLFOK DIFSAIGWIP PAOTMNALEE YKIRAOQL

```

```

351 IALCSTVPGY WETVAFIDRI GRAIQLMGF FMTVEMFAL ALPYHHWTLK
401 DNRIGFVVMY SLTEFFANEG PNATTEVVPA EIFPARLST CHGISAAAGK
451 AGAMVGAFGF LYAAOPTDPK KTDAGYPAGI GVRNSLIYLG CVNFIQLMFLT
501 FLVPESKGS LEEMSRENEG EETVAEMRA TSGRTVLEKF

!!AA_SEQUENCE 1.0
ID Q41479 PRELIMINARY; PRT; 527 AA.
AC Q41479;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER 2.
GN SPT2.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DESIREE;
RX MEDLINE=97246321; PubMed=9090882;
RA Leggewie G., Willmitzer L., Rieseneler J.W.;
RT "Two cDNA's from potato are able to complement a phosphate uptake
RT deficient yeast mutant: Identification of phosphate transporters from
RT higher plants.";
RL Plant Cell 9:381-392(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: X98891; CAA67396.1; -
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR Transmembrane.
KW SEQUENCE 527 AA; 57796 MW; 2F6D8FD9FA50C85C CRC64;
SQ

Q41479 Length: 527 May 30, 2002 09:32 Type: P Check: 2263 ..

1 MAVEDNNLQY LNALDLAKTQ LYHFTALIIA GMGFETDAYD LFSISIVTKL
51 LGRLYYTKPD LKPGTLPPA RVGLRHGAL VGTLAGOLF GCARLAKMR
101 KKVYGMIVL MIVCSVASGL SLGNTPKVYM TTLCPFRFWL GFGIGDYPPL
151 SATIMSEYAN KKTGAPLIAA VFAMOGFGIL FSGIYALITA AGFDHAYKAP
201 TFAENAVST VPQADYIWR ILMFGLPAA LTYIWRMKMP ETARYTALVA
251 KDAKRAQDM GRVLOVEIES EBAKIEQISR DETNOGLFS WEVVRHGLH
301 LFGTGSTWFL LDIAFYSONL FOKDVFSAVG WIPKAPTMA VOELYKIARA
351 QTLIALCSTV PGWFTVAFI DIIGRAIQL MGFEFTVEM FAIAIPYHHM
401 TLEANRIGFI VMYSLTEFFA NEGPNAITFV VPAELPPARL RSTCHGISAA
451 AGKAGAIVGA YGFLYAOSK DPMKTDAGYP AGIGIKNSLI VLGFINLGM
501 VCTFCVPESK GKSLEASOE TISTGEA

!!AA_SEQUENCE 1.0
ID Q82003 PRELIMINARY; PRT; 538 AA.
AC Q82003;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PUTATIVE PHOSPHATE TRANSPORTER.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

```

```

OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. EARLY MECH; TISSUE=ROOT;
RA Tahiri-Alaoui A., Avrova A., Antoniw J.F.;
RT "A phosphate transporter from tomato roots up-regulated during
RT arbuscular mycorrhiza colonization.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: Y16125; CAA76075.1; -
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR Transmembrane.
KW SEQUENCE 538 AA; 58829 MW; F93FB158A4A87115 CRC64;
SQ

```

Q82003 Length: 538 May 30, 2002 09:32 Type: P Check: 5788 ..

```

1 MANDLQVLNA LYVAKTQLYH FTAIVIAGMG FETDAYDLFC ISMTKLLGR
51 LYYHHDGALK PGLPPNVA AVNGVAFCGT LAGLFFGWL GDKMGRKKVY
101 GMTLMIMVIC SIASGLSFGH TPKGVMITLC FERFWLGFPI GDYPLSATI
151 MSEYANKKTR GAFIAAVFAM QGFGILAGM VAIIVSAAFK GAFAPAYEV
201 DAIGSTVPOA DFVWRIIMF GAIPAGLTVY WRMKMPETAR YTALVAKNLK
251 QAANDMSKVL QVEIEAPEK VTAILKQGA NDFGLFTKEF LRRHGLHLG
301 TASTWFLDI AFYSQNLFOR DIFSAIGWIP PAOTMNALEE VYKIAAQTL
351 IALCSTVPGY WETVAFIDRI GRAIQLMGF FMTVEMFAL AIPYHHWTLK
401 DNRIGFVVMY SLTEFFANEG PNATTEVVPA EIFPARLST CHGISAAAGK
451 AGAMVGAFGF LYAAOPTDPT KTDAGYPPI GVRNSLIYLG CVNFIQLMFLT
501 FLVPESNGKS LEDLSRENEG EETVAEIRA TSGRTVPY

!!AA_SEQUENCE 1.0
ID Q94DB8 PRELIMINARY; PRT; 555 AA.
AC Q94DB8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PUTATIVE PHOSPHATE TRANSPORTER.
GN P0694A04.13.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0694A04.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003294; BAB63731.1; -
DR SEQUENCE 555 AA; 60087 MW; F293412D5F6A3B76 CRC64;
SQ

```

Q94DB8 Length: 555 May 30, 2002 09:32 Type: P Check: 3975 ..

```

1 MADADGSNL AVLDAIDSAR TOMYHKAIV IAGMGFTDA YDLFCISTVS
51 KLGLRLYYOP DGSTDSKPGA LSKTANNVI GVALVGLMG QLVGFYFGDK
101 LGRKRVYGVY LILMAACAIG SGLSFGSGRK AVIGTLCFER FWLFGIGGD
151 YPLSATIMSE YSNKKTGAF IAAVFAMGV GIIFAGLVSM IVSSIFLTYN

```

201 KAPSYKGNHD LSRQMPADY VWRIVLMIGA FPAALATFYWR MKMPETARYT
251 AIIDGNAKQA ANDMQKVLST EIEAEQEKLA KFNANANNPL LSMEFARRHG
301 LHLIGTTTTW FLUDIAFYSQ NLQOKDIFPA MGLISGAIEV NALTEMFQIS
351 KASFLVALLG TFPGYWTVTA LIDKMGXYMI QLIGEFMMSM FMLAMGILYD
401 YLKTTHHFLFG LLYALTEFEFA NFGPNSTTFV LPALEFPTRV RSTCHAISAA
451 AGRAGAIVAA FGIOKLTVNS QVKSIIKALI ILSITNMLGF FTFELVPETM
501 GRSLEEISGE DGNMGAGGGG APAANAGVG VSASDVSRDE KRPAASSTEMQ
551 TSMHA
11AA_SEQUENCE 1.0
ID Q949N0 PRELIMINARY; PRT; 549 AA.
AC Q949N0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 62.4 KDA PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E., Lam B.,
RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT *Full Length cDNA of gene F13M14.22/AT3g10500 (GI:12322791).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF051003; AAK93680.1; -
KW Hypothetical protein.
SQ SEQUENCE 549 AA; 62356 MW; 89903EF6B7AFB5F5 CRC64;
O949N0 Length: 549 May 30, 2002 09:32 Type: P Check: 1025 ..
1 MGRGSVTSLA PGFRHPTDE ELVRYLLKRR ICNKPFPKDA ISVTDVYKSE
51 PWDLPDKSRL KSRDLWYFF SMLDKKRRNG SKTNKATEMG YWTTTGKDR
101 ILNGSKVVGM KKTLYVHKGR APRGERTNV MHEVRLVDQD LDKTGVHDA
151 FVLCRIFOKS GSGPKNGEQY GAFVVEEWE EEDDMTFVPD QEDLGSEDHV
201 YVHMDIDOK SENFVYVDAL PIPLNFIHG SSNNVETNYS DSINYIQOTG
251 NYMDSGGYFE QPAESYEKQD KPIIRDRDGS LQNEGIGCGV QDKHSETLOS
301 SDNIFGTDTS CYNDFPVEEN YLIGEAFLDP NSNLENDGL YLETNLDSST
351 QODGFDEFDY LTFDEFETDP SOLMGNEVDYF PDQELLFQEV ETKELEKEET
401 SRKHYVEEK EKDEASCSKQ VDADATEFEP DYKYPPLKKA SHMLGAIPAP
451 LANASEPTK DAATRLHAAQ SSGSVHVTAG MITIDSNMG WSYGKNENLD
501 LILSLGLVQG NTAPEKSGNS SAVAMLIIFMC FWVLLLSVSF KVSILVSSR
11AA_SEQUENCE 1.0
ID Q946Y1 PRELIMINARY; PRT; 520 AA.
AC Q946Y1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CERAMIDE GLUCOSYLTRANSFERASE.
OS Gossypium arboreum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX NCBI_TaxID=29729;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21424002; PubMed=11443131;
RA Leipzig M., Warnecke D., Zahringer U., Ott C., Muller F., Hube B.,
RA Heinz E.;
RT *Glucosylceramide Synthases, a Gene Family Responsible for the
RT Biosynthesis of Glucosphingolipids in Animals, Plants, and Fungi.";
RL J. Biol. Chem. 276:33621-33629(2001).
DR EMBL; AF367245; AAK73021.1; -
KW Transferase.
SQ SEQUENCE 520 AA; 58616 MW; CDB77062770BE715 CRC64;
O946Y1 Length: 520 May 30, 2002 09:32 Type: P Check: 9530 ..
1 MSAALDPVDW LLFSLSKAFR SPLAVFVQIQ GCVICITLAI GWAPAAVVRN
51 REINRMKDM KCGNSPAFLC HDINELEHTN QVNLPRVTVV MPLKGGEHN
101 LHNWKSQITS LYGPFLFLE VVESTEDPAY HAVSRILRDE KDDVDAKIIV
151 AGLSTGCSOK IHNQVGVGR MHKDKRYVLE LDDVRLHPG SIGALTAEEME
201 KNPEIFIQTG YPLDPSGSL GSYCIYEYHM PCSMGFATGG KTFELMGCM
251 MQADDERRD NYGVVSGLRD GGYSDMTLA AIAGAHKRLI TSPVAVFPH
301 PLASDISFSR YWNYLRQTF VLESYISRVN WLMNRGLFSF HCYLSWGFVA
351 PYFMAVHIA AALQIYIKGY SYEETCTTS GILLASCLAI CTLELLSMW
401 NLTRIEVQLC NMLSPAPKL SLDYYNMSLI FVALVDNFL YPISAFRSHF
451 SOSINWGSIR YHLKNGKINK IERNKGRGPK FTDLGCKHLY GKKGAPPKAS
501 FLSSIASRLC QWHQPKYEV
11AA_SEQUENCE 1.0
ID Q945E6 PRELIMINARY; PRT; 538 AA.
AC Q945E6;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN OSPT1.
OS Oryza sativa subsp. indica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39946;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. IR36; TISSUE=ETIOLATED SHOOT;
RA Kim D.-H., Nam J., Jeong S.-J.;
RT "Oryza sativa OSPT1 mRNA for phosphate transporter.";
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF416722; AAL1542.1; -
SQ SEQUENCE 538 AA; 58686 MW; 240FF749227F35A0 CRC64;
O945E6 Length: 538 May 30, 2002 09:32 Type: P Check: 7000 ..
1 MAGELAVLNA LDSAKQWYH FTAIYIAGMG FETDAYDLFS ISLVTKLLGR
51 IYFNPASKS PGSLEPNVSA AVNGVAFCGT LAGQLFFGWL GDKMGRKKVY
101 GMTLMNVIC CLASGLSFGS SAKGYMATLC FPRFWLGFGI GGDYPLSATI


```

151 MSEYANKRTR GAFIAAVEAM QGFGNLTGCI VAIIVSAEK ARPDAPAYRD
201 DRAGSTVPOA DYAMRIVIML GAIPALITYY WRMKMPETAR YTALVAKNAK
251 QAAADMTQVL NVEIVEEOEK ADEVARREOF GLFSROFLRR HGRHLGTTV
301 CMFVLIDIAFY SSNLFQKDIY TAVQWLPKAD TMSALEMEFK ISRAQTLVAL
351 CGTIPGYWFT VFFIDIIGRF VIQLGFEFM TAFMLGLAV YHMTTPGTT
401 SGFVVMYAFY EEFANFGPNS TTFIVPAEIT PARLRSTCHG ISAAAGKAGA
451 IVGSFGFLYA AOSTDASKTD AGYPPGIGVR NSLFFLAGCN VIGFFFTFLV
501 PESKGSLEE LSGENEDDD VPEAPSTADH RTAPAPPA

11AA_SEQUENCE 1.0
ID 0944R4 PRELIMINARY; PRT; 519 AA.
AC 0944R4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE AT2G19880/F6F22.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
OC "Arabidopsis cdna clones.";
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF424585; AAL1579.1; -.
SQ SEQUENCE 519 AA; 58713 MW; 72579B979D0F4BA5 CRC64;

0944R4 Length: 519 May 30, 2002 09:32 Type: P Check: 3251 ..

1 MSTLDSIDAI LFLSRAFTS PFAVFVQIOG CTICLLALG WLLAEYVRNR
51 EVKRIKNSIK AGNSLAFLYQ DINELHSRQ VKLPRVSVM PLKGFGEHNL
101 HNMRSQITSL YGPLEFLFY VESTEDPAYH AVSRLLSMYQ DHVEAKVVA
151 GLSTTCSQKI HNQLIGVEKM HKDTKYVLEL DDDVRLHPT IGALTTEMEK
201 NPEIFIOTGY PLDLPSTLG SYCIYEYHMP CSMGFATGGR TFFLWGCGMM
251 MHADFRQDR YGVVSGLRDG GYSDDMTLAS LAGAHKRLIT SPRAVAFPH
301 LASDLSFGRY WNYLRKQTEY LESYISKVM IMNKALFAVH CYLSWGEYAP
351 YVMALIHITS ALRIYIKGYH QLEDTTSASG MMLVITLAIK TFIELSMWN
401 LTRREVQLCN MLSPEAPRLS LATYNMGLVF VAMLVDNFLY PISAFRSHS
451 QSIMNSGIRY HLKDGKIEKI ERKDMGPTK TDLGKHLGY KKGAPQKASF
501 ISSLGRNLAH WRQPKKFDV

11AA_SEQUENCE 1.0
ID 0943Y7 PRELIMINARY; PRT; 125 AA.
AC 0943Y7;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

```

```

DE INORGANIC PHOSPHATE TRANSPORTER (FRAGMENT).
GN PT.
OS Thinopyrum intermedium.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticaceae; Thinopyrum.
OX NCBI_TaxID=85679;
RN [1]
RP SEQUENCE FROM N.A.
RA Davies T.G.E., Ying J., Xu Q., Li Z., Gordon-Weeks R.;
RT "Analysis of high-affinity phosphate transporter expression in alien
RT translocation lines of Chinese winter wheats.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ413964; CAC88705.1; -.
FT NON_TER 1 1
FT NON_TER 1 1
FT NON_TER 1 1
SQ SEQUENCE 125 AA; 13348 MW; BCF6E0C47CB27122 CRC64;

0943Y7 Length: 125 May 30, 2002 09:32 Type: P Check: 3376 ..

1 DLFICIALVTK LLGRITYTDP ALNEPGLTPA NVSAAVNGVA LCGTLAQGLF
51 FGWLGDKLGR KSVYGFLLIL MYLCSIASGL SFGHEAKGYM GLTCEFFRWL
101 GRGVGDYPL SATIMSEVAN KKTGR

11AA_SEQUENCE 1.0
ID 093WR0 PRELIMINARY; PRT; 525 AA.
AC 093WR0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER.
GN PT2.
OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticaceae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=PHOSPHATE STARVED ROOTS;
RC Davies T.G., Ying J., Xu Q., Li Z., Gordon-Weeks R.;
RT "Analysis of high-affinity phosphate transporter expression in alien
RT translocation lines of Chinese winter wheats.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ344240; CAC69855.1; -.
SQ SEQUENCE 525 AA; 57447 MW; E7034BDAF773F7E3 CRC64;

093WR0 Length: 525 May 30, 2002 09:32 Type: P Check: 9360 ..

1 MATEQLNVLK ALDVAKTQLY HFKAVVIAGM GFFTAYDLF CIALVTKLLG
51 RIYITDPALN EPGHLPANVS AAVNGVALCG TLAGQLFFGW LGDKLGRSV
101 YGFTLIMVL CSIASGLSLG HEAKGYMGTL CFFRFWLFGV VGGDYPLSAT
151 IMSEYANKKT RGTFAAVEFA MGFGLFEGT IVTIIVSSAF RHAFPAPFFY
201 IDAASIGPE ADYVWRIIVM FGTIPALTY YWRMKMPETA RYTALIAQNT
251 KOATSDMSKV LNKEISEENV QGERATGDTW GLFSROFMKR HGVHLATTS
301 TWFLLDVAFY SONLFQKDIK TKIGWIPPAK TMNALLELYR IARAQALIAL
351 CGTVPGYWFT VAFIDIIGRF WIQLMGFTMM TIFMLAIP YDYLVKPGHH
401 TGFVVLVGLT EEFANFGPNS TTFIVPAEIT PARLRSTCHG ISAAAGKAGA
451 IIGAFLYA SODQKPEYG YSRGIGMRNA LFLVLAGTNEL GLFLSLVPE
501 SKGKSLEELS KENVDDDTI VPTGV

```



```
11AA_SEQUENCE 1.0
ID Q93W09 PRELIMINARY; PRT; 525 AA.
AC Q93W09;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER (FRAGMENT).
GN PT2.
OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PHOSPHATE STARVED ROOTS;
RA Davies T.G., Ying J., Xu Q., Li Z., Gordon-Weeks R.;
RT "Analysis of high-affinity phosphate transporter expression in alien
RL translocation lines of Chinese winter wheats.";
DR EMBL; AJ344241; CAC69856.1; -.
FT NON_TER 1 1
FT SEQUENCE 525 AA; 57481 MW; 367463B2F9D7B46A CRC64;
SQ

Q93W09 Length: 525 May 30, 2002 09:32 Type: P Check: 9330 ..

1 MATEQNLNLK ALDVAKTQLY HFKAVVIAGM GFPTDAYDLF CIALVTKLIG
51 RIYYTDPALN EPGHLPANVS AAVNGVALCG TLAGQLFFGW LGDKLGRKSV
101 YGFTLLIMVL CSIASGLSFG HEAKGVMGTL CFFRFLGFG VGGDYPLSAT
151 IMSEYANKKT RGTFFIAVFA MCGFGILFGT IVTVSSAF RHAFPAPPFY
201 IDAASIGPE ADYVWRIIVM FGTIPALTY YWRMKMPETA RYTALTAGNT
251 KQATSDMSKV LNKISEENV QGERATGDTW GLFSRQFLKR HGVHLLATTS
301 TWFLLDVAFY SQNLFOKDIF TKIGWIPPAK TMNALEELYR IARAQALIAL
351 CGTVPGWFT VAFIDIIGRF WIQLMGFTMM TIFMLAIAIP YDYLKPGHH
401 TGFVLYGLT FFFANRGPNS TTFIVPAEIF PARLRSTCHG ISAATGKAGA
451 IIGAFGLYA SQDQKRPETG YSRGIGMRNA LFLVLAGTNFL GLLESLVPE
501 SKGKSLEELS KENVGDDDTI VPTGV

11AA_SEQUENCE 1.0
ID Q93W08 PRELIMINARY; PRT; 525 AA.
AC Q93W08;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE INORGANIC PHOSPHATE TRANSPORTER (FRAGMENT).
GN PT2.
OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PHOSPHATE STARVED ROOTS;
RA Davies T.G., Ying J., Xu Q., Li Z., Gordon-Weeks R.;
RT "Analysis of high-affinity phosphate transporter expression in alien
RL translocation lines of Chinese winter wheats.";
DR EMBL; AJ344242; CAC69857.1; -.
FT NON_TER 525 525
FT SEQUENCE 525 AA; 57377 MW; 81ECBC50849F6689 CRC64;
SQ
```

```
Q93W08 Length: 525 May 30, 2002 09:32 Type: P Check: 8781 ..

1 MATEQNLNLK ALDVAKTQLY HFKAVVIAGM GFPTDAYDLF CIALVTKLIG
51 RIYYTDPALN EPGHLPANVS AAVNGVALCG TLAGQLFFGW LGDKLGRKSV
101 YGFTLLIMVL CSIASGLSFG HEAKGVMGTL CFFRFLGFG VGGDYPLSAT
151 IMSEYANKKT RGTFFIAVFA MCGFGILFGT IVTVSSAF RHAFPAPPFY
201 IDAASIGPE ADYVWRIIVM FGTIPALTY YWRMKMPETA RYTALTAGNT
251 KQATSDMSKV LNKISEEDV QGERATGDTW GLFSRQFLKR HGVHLLATTS
301 TWFLLDVAFY SQNLFOKDIF SKIGWIPPAK TMNALEELYR IARAQALIAL
351 CGTVPGWFT VAFIDIIGRF WIQLMGFTMM TIFMLAIAIP YDYLKPGHH
401 TGFVLYGLT FFFANRGPNS TTFIVPAEIF PARLRSTCHG ISAATGKAGA
451 IIGAFGLYA SQDQKRPETG YSRGIGMRNA LFLVLAGTNFL GLLESLVPE
501 SKGKSLEELS KENVGDDDAI APTGV

11AA_SEQUENCE 1.0
ID Q9D6E1 PRELIMINARY; PRT; 179 AA.
AC Q9D6E1;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE 2900075G08RIK PROTEIN.
GN 2900075G08RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=HIPPOCAMPUS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Wittaker C., Wilmink L.,
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK013789; BAB28996.1; -.
DR MGD; MGI:1920228; 2900075G08RIK.
SQ SEQUENCE 179 AA; 21105 MW; B738709483743E82 CRC64;

Q9D6E1 Length: 179 May 30, 2002 09:32 Type: P Check: 164 ..

1 MATFPAFCY MSLVLCAL IFFAIWHIIA FDELRTDFKS PIDQNPVHA
51 RERLRNIERI CFLLRKRVGV PPGRRKGRR ERGOQLVPE YSIHSLFCIM
101 FLCAQEWLTL GLNPLFLFYH FWRYPHCPAD SSELAYDPV VMNADTLSTYC
151 QKEAWCKLAF YLSFFYYLY CMYTLVSS
```

```

11AA_SEQUENCE 1.0
ID O9CZ13 PRELIMINARY; PRT; 139 AA.
AC O9CZ13;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 2310040A02RIK PROTEIN.
RNAC OR 2310040A02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Oikido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK012585; BAB28335.1; -.
DR MGD; MGI:1913275; Rnac.
SQ SEQUENCE 139 AA; 15596 MW; 6FBFBAA4C137FAC0A CRC64;

O9CZ13 Length: 139 May 30, 2002 09:32 Type: P Check: 6690 ..

1 MATQAHSLSY AGCNFLRHWL VLSTLSGRV KIRVRARD NPGLRFEAS
51 FIRLDKRTN GSRIEINGT TLLYQPGLL YGGSVEHDCS VLRGIGYLE
101 ALLCLAPFMK HPLKIVLRGV TNDQVPSVD VLKATRSLY

11AA_SEQUENCE 1.0
ID O63363 PRELIMINARY; PRT; 475 AA.
AC O63363;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RAT NOERVTHROD ALPHA-SPECTRIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88094373; PubMed=3336352;
RA Leto T.L., Fortugno-Erikson D., Barton D.E., Yang-Feng T.L.,
RA Francke U., Harris A.S., Morrow J.S., Marchesi V.T., Benz E.J.Jr. .;
RT "Comparison of nonerythroid alpha-spectrin genes reveals strict
RT homology among diverse species.";
RL Mol. Cell. Biol. 8:1-9(1988).
DR EMBL; M19726; AAA41678.1; -.
DR HSSP; P07751; IAJ3.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00435; spectrin; 5.
DR SMART; SM00150; SPEC; 4.

```

```

FT NON_TER 1 1
SQ SEQUENCE 475 AA; 54713 MW; 3E0468C6FB6E03B7 CRC64;

O63363 Length: 475 May 30, 2002 09:32 Type: P Check: 7377 ..

1 RDMDEESWI KEKLLVGE DYGRDLTVQ NLRKKHRL ERLAEPEQAI
51 QGVLDGSKKL IDNTIGREE IQORPEQFVE HMKELQILAA ARGQRLSESL
101 EYQOFVANVE EEEAMINEKM TLVASEDYGD TLAAIQGLK KHEAFETDFT
151 VHKDRVNDVC TNGODLIKN NHHEENISSK MKGLNGKYSV LEKAAQORKA
201 KLDENSAFLQ FNNKADVES WIGEKENSLK TDDYGRDLSS VQTLTKQET
251 FDAGLQAFQO EGIANITALK DQLLAKKHQ SKAIEARHAS LMKRWTQLFA
301 NSATRRKKLL EAQSHFRKVE DFLAFAKKA SAFNSWFENA EEDLTPVRC
351 NSLEIEIKALR EAHDAFRSSL SSAQADFNOL AELDROIKSF RVASNPYTW
401 TWEALEETWR NIQIKERE LELQEQRRQ EENDKLQEF AQHANAFHOW
451 IQETRTYLLD GSCIGRRVGN SGISA

11AA_SEQUENCE 1.0
ID O9DCJ5 PRELIMINARY; PRT; 172 AA.
AC O9DCJ5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 0610033L03RIK PROTEIN (RIKEN CDNA 0610033L03 GENE).
GN 0610033L03RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Oikido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX Strausberg R.;
RA Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK002744; BAB22322.1; -.
DR EMBL; BC012416; AAH12416.1; -.
DR MGD; MGI:1915625; 0610033L03RIK.
SQ SEQUENCE 172 AA; 19992 MW; 7DF62AB86B5BF684 CRC64;

O9DCJ5 Length: 172 May 30, 2002 09:32 Type: P Check: 5548 ..

1 MPQIVELPTL EELKVEEVK SSAVLKAAH HYGAQCCKTN KEFMLCRWEE

```

```

51 KDPRLCKEG KLVNGCALNF FROIKSHCAE PFTEYWTCLD YSNMQLFRHC
101 RQOQAKFDOC VLDKLGWVRP DLGOLSKVTK VKTRDPLPEN PYHSRARPEP
151 NPVIEGDLKP AKHGTRFEFW TV

!!AA_SEQUENCE 1.0
ID 091YD4 PRELIMINARY; PRT; 1507 AA.
AC 091YD4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TRANSLANT RECEPTOR POTENTIAL CHANNEL 7 (FRAGMENT).
GN TRP7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA kashuba V.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ344343; CAC69081.1; -
KW Receptor.
FT NON_TER 1507 1507
FT SEQUENCE 1507 AA; 172364 MW; 86E552DE25939785 CRC64;
SQ 091YD4 Length: 1507 May 30, 2002 09:32 Type: P Check: 1512 ..

```

```

1051 QEVOEHTDOI WKFORHDLIE EYHGRPPAP PLLLSHLQL LIRIVLKIP
1101 AKRHQLKNK LEKNEETALL SWELYKENY LONQOYQOK RPEQKIODIS
1151 EKVDIMVDLL DMDQYKRSGS TEORLASLEE QYQVTRALH WIVTTLKDSG
1201 FGSAGALLTL APQRAFDEPD AELSIRKVE EPGDGYHSA RHLLYPNARI
1251 MRFVPNEKV PMAAEFLIYD PPEYTAEKDV ALTDVPGDTA EPLSKISYNV
1301 VDGPTDRSF HGYYVEYGF PLNDMGRGTL RGRGSLSWFG PNHTLQPVVT
1351 RMKRNQGGAI CRKSVRKMLE VLVMKLPRSE HNALPGSRE PGEMLPKRLK
1401 RVLROEFVVA FETLLMQGTE VYKGYVDDPR NTDNAMIETV AVSIHFQDON
1451 DMELKRLLEN LHTHDPKELT RDLKLTSEWQ VVDRRIPLYA NHKTILOKVA
1501 SLFGAHP

!!AA_SEQUENCE 1.0
ID 056074 PRELIMINARY; PRT; 3005 AA.
AC 056074;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE GENOME POLYPROTEIN.
OS Hepatitis GB virus A.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
OX NCBI_TaxID=39112;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97126113; PubMed=8971037;
RA Leary T.P., Desai S.M., Yamaguchi J., Chalmers M.L., Schlauder G.G.,
RA Dawson G.J., Mushahwar I.K.;
RT "Species-specific variants of GB virus A in captive monkeys.";
RL J. Virol. 70:9028-9030(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97437486; PubMed=9292019;
RA Leary T.P., Desai S.M., Erker J.C., Mushahwar I.K.;
RT "The sequence and genomic organization of a GB virus A variant
RT isolated from captive tamarins.";
RL J. Gen. Virol. 78:2307-2313(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98120818; PubMed=9460920;
RA Erker J.C., Desai S.M., Leary T.P., Chalmers M.L., Montes C.C.,
RA Mushahwar I.K.;
RT "Genomic analysis of two GB virus A variants isolated from captive
RT monkeys.";
RL J. Gen. Virol. 79:41-45(1998).
RN [4]
RP SEQUENCE FROM N.A.
RA Erker J.C., Desai S.M., Leary T.P., Chalmers M.L., Montes C.C.,
RA Mushahwar I.K.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF023425; AAC40502.1; -
DR HSP; P27958; 1HEI.
DR MEROPS; S29.002; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002518; HCV_NS2.
DR InterPro; IPR004109; HCV_NS3.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRp.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR001005; Myb_DNA_bind.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.

```

DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; HCV_RdRP; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
SQ SEQUENCE 3005 AA; 323793 MW; BBA07D1ADFAB81C2 CRC64;

056074 length: 3005 May 30, 2002 09:32 Type: P Check: 8233 ..

1 MEVLLVLLK TALAGAFLOP ASHACHAAGY YLTNCCSDS EITACFDGCG
51 LVSVCCTVCD GRCWDLYRPG VATRPGHPCG ELFGALMSTG SPAVVAAYVA
101 GITGLGEPFS AALLAALAR VEMTPRPNT CLMCDLAIQ SEAMDIVEIM
151 RDTFTVKWV MTLPRMLWG LVGMSFALMV VVLLLEQR VVMVLIVLAM
201 AGSVEFTGG CRCHVNGSIV AIPGAKPEDR PNATQVCYCP FGQMYWLPAL
251 CAGLAWRNGD VSGTTRDLPV RCPESMHGRG SVMCWGSAH MSWLATEYR
301 LMEQIPGSAL CHEFAVGTTD RPHPATDVLG THGIPCASCY VDRRASWCGN
351 CTRDCWEKTEG NKRLSEACG LGSRLTAELW AHVVDGTEG KIVTDGERP
401 RYPSHGVGT HHTAVVAARN YTVSDIGGYW HAACPNPSS PSALPKLIPG
451 RPYNACLTIC KEGRLNTAW QAPGXFAPL FTECNWPRRS GVRVCEGYAF
501 DFPGSKTGFI RSRAGFOQIA RAGSMPPHRW LITDYLFILL VLMKLAEARL
551 VPLVAVALYW WFNMAEAAT LRILHPVNI TTAAPMTWPA LPVPTVCPT
601 KAGAVGDALY MAACLAGNAT LGVSSVWVG AYGAEEAGAR GLMRWNGER
651 SFCAGLAWLT NVGAYLPVVE AAVGPEFVSA PLLVMAMEDS IAVFILMAN
701 IWAYMSDHLG VKLAGLVAAH LAKGALPLVL LVAAXVTRHR HSVLGFEVCY
751 SLDADASPSD WSMWALAAYV SWCLLTIGLT THGRLRKLK FYSYWCREFH
801 AVRMRLLCSK VGRGRDWRV TAVVVAAGII FPREVVRCSA ALTALAALD
851 SIDYLLETLI LTAAPARAA RLDDSLTFLG DADLTRAFVR RLRRGVTLF
901 QHCGQVSXGA AAILXDLGVA LEPVSVTARD CYIVRDAART LACGQVEGL
951 PVVARRGDEV LVGVFPVSRA LPPGFVPTAP VVMQRGIGF FSVVKTSMLG
1001 RDERHEGSI VVLGTSTTRS MGTGVNGVMY TTFHGSNART LAGPVGPVNC
1051 RWWSPSDOVA YVPLPSGASC LEPCKCGTOS VMCIRNDGAL CHGRLSKVE
1101 LDLPTEISDF RGSSGSPILC DEGHVVGMMY SVLHRGVAVT GVRVYKPMET
1151 LPKDSQVKSX APPVPGKTGF TEAPLYLPTG SGKSTRIPLE YTKAGHKVLY
1201 LNPSIATVRA MGPMYKLSG QHPSIYCGHD TTAYSRTGGS PLTYCTYGRF
1251 MANPRRYLRG ADIVICDECH VTDPITSVLGM GRARLLAREC GVRLLZATA
1301 TPPGAPLAQH ESIKEVPLGV DGEVAFYGHK LPVERYPYGR HLLFCHSKVE
1351 CNRLHAALST AGCNAVVIYR GNEQEIPIAGD VCVCATDALS TGYTGEPSTV
1401 TDCGLMVEEV VEVTLDPTIT ISVRTTPAPA ELRAQRGRG GRGSQGTYYY
1451 AMTASAPAGT LRSGLPMAAV EAGVAVMYNE PDMTADILRA YDACPYYTAI
1501 TASVGEAINF FSGLVPMRNY POVAMAKSHQ HNWPLLVGVQ RTMQQEAANYA
1551 GPGDGPEWAG LAGTGPIPLL CRWGARPSS VAPHHWVDL QARLVAEGY

1601 SPCYAGPILL VGLALAGAV LAHWTGSLVV VTSWRVNGG NPLIQGSTRG
1651 VSTSAPYQLT VCVEGEQTPA DGKCAAEAVQ LXESTCGWGP MAASFDCAGM
1701 KGVLDSMRTT AAAAVEKSDS LMRSCANNY CPPGGATSA SAFASLDTK
1751 FAQAWDAIFT NGRSLVGLV AGYGARRNP LGVAAAFIMG MSAGHQVHR
1801 LAAALLGVG GTMLGTPSVG LAMGAYFAG GSITSSWLSA IVAVLGWEG
1851 AXNAASTFD FLTGRAELKD MFLVSCXAS PGASVAGVAL GLLWSMKKG
1901 VGEDWVNRII TLIPRGSVLP DGEFFKSEFT ERVSTILRKM SLSRWMTLV
1951 ERRELDLETP CSSMLMDLID WLVRFGRYIG RRLKGMPSV RVLVGCYTPG
2001 WGSWVGEGH IEARCACGCI ITADVEEKL VDVHYSSRLC SNYLGITPVY
2051 SAAGSDAEP EVPAGPALYQ IGVAEWQLV RKDKTLVGA SSVYHLIDE
2101 LRRAIRGPPM FVGGVGSWE APLQOPPLYV RAGQSVREDD VRYSLPHTLA
2151 LPPPRPPPP LAMP PPPPP EQWTEEEEA DLREARARAI EAVNERLBP
2201 NPEAAQALD ALEAAVSLI PHVGAILGDD CSCEESFGH FIPEDAVDY
2251 PIGMEVQVG PLRDQARDLG DRLAVLGARL ESLAEAHPEA SLNTERRMG
2301 ELIDTLADVQ ARLDVTCRSD TSSSSFEQIS LSDSEPTIV EGGLKLEYVR
2351 POPVRFKDLI RPEGAKRLV TVROSCADR SATRAFALSL PIAAVTATLS
2401 FDLTDHTVSD STGRVLDPLE LLQNAVGLI VACRSESSV SYSYIMSGAP
2451 LNTGRHQPAP MTRPIGTHIT ADTTKVYYTD PNOGERAAK VTIMRGSRY
2501 DAHYRGVSE VLQAKTIKS PGWYDEAIA KVSRAAAGF GSKVTGNNMT
2551 TPAARAEVNN MLAKIKTKQE VPFTLVTKRE VEFQKTRKP PRFICFPBLD
2601 FRIAEKMILG DPLVAKGIL GKSYLQYTP NQVKLMVDL WRQKHBRAT
2651 TVDATCFDSS IDERDMAVET EVFAASPDP DLVRALGSY AEGPMVSPKG
2701 VPLGVRKCRS SGVLTSSAN SITCYIKVKA AARAVGLVDP DFLIAGDDCV
2751 IYEDGEDH ADALRVALGN YGYDCKPKTH ASLDTAESCS SYLAECNVGT
2801 ERVWMLSTDM RKPLARAASE YSDPVSSALG TILMPWHPD VRWVLLPHIL
2851 IMAFRGGTP DDLVCEVQG NHYSPLRYL PEVLVSLHGP RCLRVADST
2901 KTKMEAGAAI RDLGMHTLAF YRKRAGNVRT RLLRGKGWG RLARALLMHP
2951 GLKEHPPSIK SIPEFKMATP YEHHETWWS GEKPPWYMDV KCLFGLYCAF
3001 LTALL

!!AA_SEQUENCE 1.0
ID Q9DW56 PRELIMINARY; PRT; 655 AA.
AC Q9DW56;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE PR139.
GN R139.
OS Rat cytomegalovirus (strain Maastricht).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_Taxid=79700;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAASTRICHT;

RX MEDLINE=20366325; PubMed=10906222;
RA Vink C., Beuken E., Bruggeman C.A.;
RT "Complete DNA sequence of the rat cytomegalovirus genome.";
RL J. Virol. 74:7656-7665(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MASTRICHT;
RX MEDLINE=20473137; PubMed=11018281;
RA Gruijthuisen Y.K., Beuken E., Bruggeman C.A., Vink C.;
RT "Rat cytomegalovirus R89 is a highly conserved gene which expresses a
RT spliced transcript";
RL Virus Res. 69:119-130(2000).
DR EMBL: AF232689; AAF99234.1; -
DR InterPro: IPR003360; US22.
DR Pfam: PF02393; US22; 1.
SQ SEQUENCE 655 AA; 73553 MW; 652CEF38F0664B47 CRC64;
O9DW56 Length: 655 May 30, 2002 09:32 Type: P Check: 4069 ..
1 MMSLRGALRR DVREPLGREHH EMPTEVWQDP TALEPGTFDE DFGCVDDDL
51 EKHIPLSRSP VYPPDDSLIK GVLGLSMSEE GRERGTRSGP LTGHDNMAAL
101 NEPDDEIAQT VSVRESREAS GHRHGGSSDS DLGSDSGRSD NEIEIVOGHS
151 WKSTALTVSK EIVEHEKKN LSRREFATVL RDEFTLYVAL SNPNEIPERI
201 RQLSGRALTL RPEHMYLLP QPMNEIRVLF GVDLRAYICC DVAVARDEL
251 IPLGACAAARS PDLHYRETFC AILMDTEGRF FLYDAETEGE FLAENLVEL
301 AHKGISACEP AYRDGATIS LPRPRTAVKK ILSACVIGLE SVNAAVTSVR
351 GTAIVLCDBA SGERILQIF DASELRKKPP FSELDDESCG AMMDYVEFRL
401 AEEWVYVGGI GTYDNGSGFV FMVDIVVLG VTGAVYGGFV EENDYVRLAD
451 DLSTLLRRGI SPEPNRFRD ALGELRLERR ALCPHEREGT RGDREPVVRQ
501 VARDLRGWL RMQLRTGPD DRCVOFAHLE EAKRQFRHPV SLEPVCVVTG
551 ESYQPRNESE QGYGVLAEGR RNFRCPRTRP DPTDERRRQ RELYQRMCP
601 PGPEVNPPEE TYTGVAIERA ARLALQKSG TSLRLPAISR AATRVSRGS
651 PMSDR
11AA_SEQUENCE 1.0
ID Q83896 PRELIMINARY; PRT; 595 AA.
AC Q83896;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TERMINAL PROTEIN.
OS Ovine adenovirus type 7.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Adenovirus.
OX NCBI_TaxID=114430;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OAV287;
RX MEDLINE=95297141; PubMed=7778275;
RA Vratil S., Boyle D., Koehrhans R., Both G.W.;
RT "Sequence of ovine adenovirus homologs for 100K hexon assembly, 33K,
RT pVIII, and fiber genes: early region E3 is not in the expected
RT location.";
RL Virology 209:400-408(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OAV287;
RX MEDLINE=97080497; PubMed=8921842;
RA Vratil S., Brookes D.E., Boyle D.B., Both G.W.;
RT "Nucleotide sequence of ovine adenovirus tripartite leader sequence
RT and homologues of the IVa2, DNA polymerase and terminal proteins.";

RL Gene 177:35-41(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=OAV287;
RX MEDLINE=97271311; PubMed=9126262;
RA Xu Z.Z., Hyatt A., Boyle D.B., Both G.W.;
RT "Construction of ovine adenovirus recombinants by gene insertion or
RT deletion of related terminal region sequences.";
RL Virology 230:62-71(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=OAV287;
RX MEDLINE=98277079; PubMed=9614874;
RA Khatri A., Both G.W.;
RT "Identification of transcripts and promoter regions of ovine
RT adenovirus OAV287.";
RL Virology 245:128-141(1998).
DR EMBL: U40839; AAD45952.1; -
DR InterPro: IPR003391; Adeno_terminal.
DR Pfam: PF02459; Adeno_terminal; 1.
SQ SEQUENCE 595 AA; 69781 MW; 026ADF6563066206 CRC64;
Q83896 Length: 595 May 30, 2002 09:32 Type: P Check: 5502 ..
1 MTARQIHQMQ QLTGQSQHTL RYLRLTLDLN NRSIARSRTI HERGITKANR
51 FENYPTVQLL DLPRGPVTS QPFEEGPPP NLLIGFYVA RVLNQLYLDQ
101 RTYSNISYKL YLSPISFERR MTQQLTDCS YSINTGYSYR AIQNVENFSQ
151 TISQIQNAV LMDRIIGSLQ ADMQFGSAI SAQONRSFO QHYQFESQAF
201 MVLNLQERDT YLIKLIKIK KALCKFLILS NMHNSCIL DLPFSEFIEL
251 FIDFSQLLEI PEARNRINK DLATVLTGK GGMHGALTL RSGTRVGLPF
301 RLPRRENTRA ITEIMRRSRG DVIRRFIDRL PVNRQAVRN VEETSSPLT
351 VPEIEDEGS DREPPSSDLS REEFNDEVIA STVDLIONLE EELTPEARRS
401 NEFNYSSEFF QLLIRFYNEN RLTDPIQKW LVYFFILEHV ASTLYLYLIN
451 LVQNRLAARN IGIQVQIIL RGRNENGEDI FTRVFNRET QAFRLYSRI
501 TRDFLGITFA NERSYSFSTP EERDQLADI DFVEDSGSIE EVINQVNTDF
551 SLDLSVEIAF RIKLSGIQVY STNEVVLRSF ERVREALNR WLQRO
11AA_SEQUENCE 1.0
ID Q91GE6 PRELIMINARY; PRT; 243 AA.
AC Q91GE6;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 26.9 KDA PROTEIN.
OS Epiphyas postvittana nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OX NCBI_TaxID=70600;
RN [1]
RP SEQUENCE FROM N.A.
RA Hyink O., Dellow R.A., Olsen M., Caradoc-Davies K.M.B., Drake K.,
RA Ward V.K.;
RT "The complete sequence of the Epiphyas postvittana
RT nucleopolyhedrovirus genome.";
RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY043265; AAK85672.1; -
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 26869 MW; B5203A27C88CAEA0 CRC64;
O91GE6 Length: 243 May 30, 2002 09:32 Type: P Check: 8973 ..
1 MGLFNQFSQY TRLPAAPQIS LAVMSYVNTV LCAYGAIVAA YLSTATSFVE

```

51 LQFLEYWML SLLINGLINV TLFROSKE AHEIYVELKM LHMYFSSAL
101 VNLAIDTAQ SASAVLVNN LIHCLALFLL FVELTVLGH TLGTANYRY
151 TKACYLVVLL VTAAVTIIIL TAENVKSSPL CNDLIMASFL TAAFVIAVY
201 WAWRKEAAGS VLQRVQLTSL YDPPPSFTNV KMEDMLKNQ MEI

11AA_SEQUENCE 1.0
ID Q9W7A2 PRELIMINARY; PRT; 380 AA.
AC Q9W7A2;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FAST TROPONIN T ISOFORM.
GN TMTF.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99292729; PubMed=10364205;
RA Bucher E.A., Dhoot G.K., Emerson M.M., Ober M., Emerson C.P. Jr.;
RT "Structure and evolution of the alternatively spliced fast tropoin T
RT isoform gene.";
RL J. Biol. Chem. 274:17661-17670(1999).
DR EMBL; AF139128; AAD42799.1; -
DR InterPro; IPR001978; Tropoin.
DR Pfam; PF00992; Tropoin; 1.
SQ SEQUENCE 380 AA; 43814 MW; 000A930E3D3689DC CRC64;

Q9W7A2 Length: 380 May 30, 2002 09:32 Type: P Check: 3716 ..

1 MSDTEVEHG EEEYEEEEEV QEEVHEPAP PPEAPEEEE KPRIKLTAPK
51 IPEGKVDFD DIQKRQNKD LIELQALIDS HFEARRKEEE ELVALKERIE
101 KRRAERAEOQ RIRAEKEKER QARLAEEKAR REEDAKRKA EDDLKKKKAL
151 SSMGASYSSY LAKADQKRGK KOTARETKKK VLAERRKPLN IDLINEDKLR
201 DKAKELMDWL YQLQTEKYDF TEQIKRKYYE ILLRCRLQE LSKLSISGT
251 GLIRLRNTAR ROEPKARLAG AGSKEPGRMA LSDVSKPCWS PSSSFTNYLC
301 SCASAITKIA TSACVGYCCC FSSFWRGIYA SSVKRKQVPC GTGETPCWHL
351 QFLSMSFIL RTPALCPVD CCVQIPGFC

11AA_SEQUENCE 1.0
ID Q90ZD4 PRELIMINARY; PRT; 491 AA.
AC Q90ZD4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INTERFERON REGULATORY FACTOR-3.
GN IRF-3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA May D.L., Grant C.E., Deeley R.G.;
RT "Cloning and Promoter Analysis of the Chicken Interferon Regulatory
RT Factor-3 Gene.";
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268079; AAK58583.1; -
SQ SEQUENCE 491 AA; 54440 MW; 90E0C80A2624B1B4 CRC64;

```

```

Q90ZD4 Length: 491 May 30, 2002 09:32 Type: P Check: 3654 ..

1 MALDSEGDG QKURFGPWL NAVSSGLYRG LCWIDPDRI FRIPWKHNR
51 KDVTSSDVEI FRAMAKASGR YEGNAEDPAK WKTNFRCALR STMFMLED
101 RSCNDDPHK YVAVASGVPN DRSGGVPAG ALQQQPOLL NHHDLALENT
151 PFDSTEGVAA AALTQVDLDL LOSVLQHCNI SALGSQPTLW AHTGDALPED
201 ALLLPQDGC LPPQFQDWR QLEPPLIGN QPLTGGCGQ DGAGALPVSE
251 ECAIPAPSPA EELLFQSANP APPPPAGDTG GLPPLDITI YYRGKMYQE
301 QYDDSRCVLA YQPLDPAVAE QRLVFPSPA SLPPRQRRY TENLLEVAGL
351 RLEORAGOLL ATRLKKCKVF WALSQOLEGG EPPLNLHRD QETTIFDERV
401 FCTELRDFRD SRERSPDFI IFLCFGQCFS STPKESKLI LKLVPOFCE
451 YWEQVQORG ASSLSGNVS LQSDSFNLF ELIEQYHMOT D

```

```

11AA_SEQUENCE 1.0
ID 025915 PRELIMINARY; PRT; 417 AA.
AC 025915;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE COMPETENCE LOCUS E (COME3).
GN HP1361.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Colton M.D., Weidman J.M., Fujii C., Bowman C., Wathley L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RT Nature 388:539-547(1997).
RL EMBL; AE000636; AAD08401.1; -
DR TIGR; HP1361; -
DR KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 417 AA; 47984 MW; B9D557BE7F7ED828 CRC64;

025915 Length: 417 May 30, 2002 09:32 Type: P Check: 8156 ..

1 MCGVFLSLLL AINLYLEYLN YQKLDSEKPT SLAQILLQY PTKDQKTYF
51 VLKLSKNMI FYTTIKEPLK NLQYRHAQFF GKIKPCSFLE SLKSCFFQTY
101 SPSLTRKQDF KSHWRHFIDS AHENALVGNL YRALFIGDSL NKDLDRANA
151 LGINHLAIS GFHLGILSVS VYFLSSLEYT PLQKRYFPYR NAFYDIGLV
201 WYFLLGYLLL LDFLPSFFRA FLMLGLGFLA CFFGVRLSF KLILACCTA
251 IALLPKLFS VGFLISVCGV WYIFLFLKHT QIFFKTSSFL MRSFOAISLS
301 ALVFLMLTI VHAFFPMFSP YQLFSIPLGL IFIVFFPLSL FLHAVGLGSL
351 LDRLLSMPLT IPTISVPSP LILGVHLFT ILSARFFKYV LSMNVLSAGF

```



```

401  FLVCCYQYII MPSLIVG

11AA_SEQUENCE 1.0
ID  P72685      PRELIMINARY;      PRT;      290 AA.
AC  P72685;
DT  01-FEB-1997 (TReMBLrel. 02, Created)
DT  01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DE  01-JUN-2001 (TReMBLrel. 17, Last annotation update)
GN  SYNTHETICAL 32.1 KDA PROTEIN.
OS  Synechocystis sp. (strain PCC 6803).
OC  Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX  NCBI_TaxID=1148;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=97061201; PubMed=8905231;
RA  Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA  Miyajima N., Hiroseawa M., Suglura M., Sasamoto S., Kimura T.,
RA  Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA  Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA  Tabata S.;
RT  "Sequence analysis of the genome of the unicellular cyanobacterium
RT  Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT  entire genome and assignment of potential protein-coding regions.";
RL  DNA Res. 3:109-136(1996).
DR  EMBL; D90899; BAA16692.1; -.
DR  InterPro; IPR003675; Ab1.
DR  Pfam; PF02517; Ab1. 1.
KW  Hypothetical protein; Complete proteome.
SEQUENCE 290 AA; 32091 MW; AD481D2FA2AEB4DD CRC64;

P72685 Length: 290 May 30, 2002 09:32 Type: P Check: 1310 ..

1  MKSTVNENF NOQAIAQLTA PWKIGVFTVM LAAAWLPFLV PLSLAIDAN
51  LRSIVVMGIL FLIFLVLLIF WSHWCYQTP LSKAYGVYGL GNNRRQGVEL
101  LRGLGLGFSF TFGFLTIQGL LGWAVLAPAG DRLMTIIMOG SLTGLGVALA
151  EELFFRGWLL KELEQGYGNK TSLASNAIIF AVLHFLKPLG EVIRTLPOFP
201  ALVVLGLSLG ITKRRHGDR L GHSIGLHGGM VWAYIIVNVG QLVTTYEKM
251  AMVTGIDRNP LSGWGIAGL CLLMLVNOG EKPWLKRLGF

11AA_SEQUENCE 1.0
ID  Q9PAP6      PRELIMINARY;      PRT;      544 AA.
AC  Q9PAP6;
DT  01-OCT-2000 (TReMBLrel. 15, Created)
DT  01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE  01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE  HYPOTHETICAL PROTEIN XF2449.
GN  XF2449.
OS  Xylella fastidiosa.
OC  Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC  Xylella.
OX  NCBI_TaxID=2371;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=20365717; PubMed=10910347;
RA  Simpson A.J.G., Relnach F.C., Arruda P., Abreu F.A., Acencio M.,
RA  Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA  Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA  Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA  Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA  Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA  Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA  Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,
RA  Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA  Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA  Krieger J.E., Kurnae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

```

```

RA  Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA  Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA  Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA  Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA  Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA  Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA  de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA  Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA  Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA  da Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA  da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA  da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
RA  de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H.,
RA  Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA  Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT  "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL  Nature 406:151-159(2000).
DR  EMBL; AE004053; AAF85248.1; -.
KW  Hypothetical protein; Complete proteome.
SEQUENCE 544 AA; 62881 MW; 45D840B178BBA046 CRC64;

Q9PAP6 Length: 544 May 30, 2002 09:32 Type: P Check: 144 ..

1  MICRRVYLV D SMNLRSSHW FVFLSVAA L LQGCQKQDSP SAKRQIIMVE
51  AKLPPRLKP DLPLPLDIE RRTFQFFWDT TNEVNGLTPTD RYPSRPEASI
101  ASVGALTAY PIGIENGWIS RNOAIDRTLT TLKFLRDAPM GPQRTGCAGY
151  KGFFYHFLDM QHGHRYSWV ELSSVDTALL MMGVLFVESY YDGEDVREKE
201  IRLIADKLYR RVDWLYLQOR KPLISMGWYP ERGFIEHDM GYNEAMVYL
251  LALGSPTHPL EPVSWDEWTR TYNKDGVETQ GQEYLAFGPL FGHQYTHWV
301  DFRDIQDQYM RERGIDYFLN SRAVLAHRD YAIIDNPMKKK DYGENWGLT
351  ASDGPQNTTQ EYRGEOQRF HYRSRGAGLF EFFDGTIAP TATVASIVEA
401  PEVVIPATLE MHRKYGDFLY SSYGFLDAFN PSFDYVPLK TGRLIPGRGW
451  VASDYIGIDQ GILLSMIANY RNDFVNVWVK KKKYVRTGLE RAGFLIGMWLQ
501  EDEVLRGPQK DERAANVSL GIAESHAPPA QTORVPSSLG KKPE

11AA_SEQUENCE 1.0
ID  Q9KVA4      PRELIMINARY;      PRT;      463 AA.
AC  Q9KVA4;
DT  01-OCT-2000 (TReMBLrel. 15, Created)
DT  01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE  01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE  RRBG PROTEIN.
GN  VC0245.
OS  Vibrrio cholerae.
OC  Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OC  NCBI_TaxID=666;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=EL TOR N16961 / SEROTYPE O1;
RX  MEDLINE=20406833; PubMed=10952301;
RA  Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA  Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA  Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA  Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA  McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA  Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA  Fraser C.M.;
RT  "DNA sequence of both chromosomes of the cholera pathogen Vibrrio
RT  cholerae.";
RL  Nature 406:477-483(2000).
DR  EMBL; AE004113; AAF93421.1; -.
DR  TIGR; VC0245; -.
DR  InterPro; IPR002047; AKH.

```


DR PROSITE; PS00256; AKH; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 463 AA; 53157 MW; 7B93DAB71BDBECFA CRC64;
Q9KVA4 Length: 463 May 30, 2002 09:32 Type: P Check: 1939 ..
1 MNVMTKKKL VVWVPLPPNS SWRGGIAQT IENIVRNISP ERKIEIVSS
51 KHAEMLVGLE KSNPNISVLT LGFRGKSTYK TIGYVSLNEV EKDSLMDIVI
101 AKLPIIPAIF RKGVMYVSQ L EYLLSLIYS HLQRRGRSS NNCRWLPTP
151 IIPYTHLGG EKFSFWDPF VFEYNKEPPL TAEYFKKLS KHFSNASAI
201 TQSRANKDY L ETVMGIESK INVIYNGSPD YSEFKQGSN LSFSEVMSK
251 EFSGASKKAA FEALVNHQ L N FSVLWRLTK NKVSNRKIVL ISTQNRPYKG
301 FDQLEVLIN LCLRRDNYDF IFTCNVPTKL KERYPSLYER IHEVTRVDNY
351 LHASLYIMSD IVLHPSNVEG GLGAYPQYEA SSVGKPSLIN TGRHVNEAE
401 EGFVDVLSS NFVNTKETVD KIEKLINSEE YMRONIDAIN RLKISWKESA
451 SNYENVEFGN ENA
11AA_SEQUENCE 1.0
ID Q9KQ57 PRELIMINARY; PRT; 118 AA.
AC Q9KQ57;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN VC2147.
GN VC2147.
OS Vibrrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_Taxid=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uterback T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004287; AAF95292.1; -.
DR TIGR; VC2147; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 118 AA; 13492 MW; B10702A234690953 CRC64;
Q9KQ57 Length: 118 May 30, 2002 09:32 Type: P Check: 2237 ..
1 MKNTTLPSE QTOQALRIA KATQPGQTK EQTKLTQGI EKGIATYKKQ
51 QKEKHROADK LRKKALKAKQ SSTEIHEAD DYAAELGDDS HTNQAKLAWY
101 LLALSWLGFI SYLWQNS
11AA_SEQUENCE 1.0
ID Q9KE45 PRELIMINARY; PRT; 155 AA.
AC Q9KE45;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE BH1013 PROTEIN.
GN BH1013.
OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_Taxid=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
RT halodurans and genomic sequence comparison with *Bacillus subtilis*.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001510; BAB04732.1; -.
KW Complete proteome.
SQ SEQUENCE 155 AA; 17487 MW; FA40F2DD6840D0A6 CRC64;
Q9KE45 Length: 155 May 30, 2002 09:32 Type: P Check: 7616 ..
1 MNTLCIIPCG NKIWDKQED IGPVAASEAY IGTLHRLCAM YAEFFDHWY
51 ILSAKHGFLF PDDVDGPDY VFSFTHSGV ITFDQTEQV HEKKLDVPH
101 VVLTGKKYR PIVEACFPPEA TVECPPLSYS GIGYMQQALK RAIEQKPLH
151 SAKRN

11AA_SEQUENCE 1.0
ID Q9KB88 PRELIMINARY; PRT; 818 AA.
AC Q9KB88;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE BH2040 PROTEIN.
GN BH2040.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_Taxid=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
RT halodurans and genomic sequence comparison with *Bacillus subtilis*.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001514; BAB05759.1; -.
KW Complete proteome.
SQ SEQUENCE 818 AA; 95517 MW; 0CD750C3815978D0 CRC64;
Q9KB88 Length: 818 May 30, 2002 09:32 Type: P Check: 9494 ..
1 MTKFINEKL NDPDQGEWC FEPHMLRSA GEPKWLQF TFTDTMKLVE
51 EKRLDEKWN ETISECNNE L EKMLDLTTKQ KKLISQLRR YKIPTAKGST
101 ELLANLGAGF QIKLEDMIEK MNYLEAQISK TQTELDRO R HHLKQVYLEN
151 ESLQEAFLS SPDAYHNGLK HPLSKLEQS RPRGQROKEQ LAAIYLQRF
201 SKNETASYYG PTNYGTFTSQ PGLLELHVNG PIKRLFMST WAVQSLANYM
251 AEDDSILPYL KPKLSPFIKR ENSNLROSSN GKLILPEIY HKMIRYSDGN
301 HNIHQIAAML DLSTAECIAR IQKLEKRII LLEIDFPTS V FDPFEKLEK
351 IOOLPNDCCS KSKWLHVINE WSSHLITWKE SESFOQRRI L TLLSENSFEQ
401 NVGLSSRKDR GKHYTDRLIV FEEARGDVE T CRIGLMQOK WKDQLEPIFK

```

451 LISGRAVEH RALTEVAYKE YQSPNSDVNF LTFALHMQGC KTEALQYAKE
501 KLKEDNRLS EWTALHPEEK GSIYAKEDV DAFCLNFPQP DVALFSPDIM
551 IAADVDAIN EGKYLILGE VHSIGIOWSV LNSYRPDQLR LNEIYHHLG
601 PTLQSWLEH VGPRAPGKTF RPELSQGTIV ENLGRSMKSR EYVRSIAELG
651 LIYENERFYV TEGEKPKLMD LETDVEPLNQ IFSLPYSKSF SIOGEHTPR
701 IEINGVFOR ERWTFACKEL LVRVNGYTDW LLDWATDLR KNYNMPRYVY
751 ARGNNPEKPI FVDENFEFTL EVLYQLLKRN EHVSIEMVP AQDSLMPTRG
801 ENKHTAEFRF SVIHENLT

11AA_SEQUENCE 1.0
ID Q9K650 PRELIMINARY; PRT; 154 AA.
AC Q9K650;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE PUTATIVE RNA POLYMERASE SIGMA FACTOR.
GN BH382.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY. ECF SUBFAMILY.
DR EMBL; AP001520; BAB07601.1; -.
DR InterPro; IPR000838; Sigma70_ECF.
DR Pfam; PF00776; Sigma70_ECF; 1.
KW Complete proteome; DNA-binding; DNA-directed RNA polymerase;
KW Sigma factor; Transcription regulation.
SQ SEQUENCE 154 AA; 18291 MW; A4FC9868ED248A03 CRC64;

Q9K650 Length: 154 May 30, 2002 09:32 Type: P Check: 5308 ..

1 MKVSELYEEL KQDIYFART ISRNEQAND LIDAMEKSL RCTELLDMPR
51 HKQKAMFYRV MKNQOLIDIR KKRRTDWD SVIPFTTAG TSYEMVDLL
101 SKLPASLSDL VFKEHFWLGLT SQEISEKGV PASVYRYRIR LAVKKLREYM
151 KEEG

11AA_SEQUENCE 1.0
ID Q9JY55 PRELIMINARY; PRT; 167 AA.
AC Q9JY55;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN NMB1733.
GN NMB1733.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MC58 / SEROGROUP B;

```

```

RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwin M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AE002523; AAF42078.1; -.
DR TIGR; NMB1733; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 167 AA; 19854 MW; 173631A183372CDE CRC64;

Q9JY55 Length: 167 May 30, 2002 09:32 Type: P Check: 3936 ..

1 MDAFDFEMQG NEKKIATPQG LDRKYFTFQY IVCNLSAISS FCLMKIHTYV
51 TILYILIAVI GIINGFLRS KVLSTVKNKQ KFLSDMPELE MPFFVEYKI
101 PILARFLNRG MKWLYRPSFF CLVAICFSIV ILNEGGWMP YFACYWGLGC
151 YPVSTYVVEL SRDYKHI

11AA_SEQUENCE 1.0
ID Q9JVL8 PRELIMINARY; PRT; 173 AA.
AC Q9JVL8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE POSSIBLE MEMBRANE PROTEIN.
GN NMA0784.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajadream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162754; CAB84067.1; -.
KW Complete proteome.
SQ SEQUENCE 173 AA; 20841 MW; 8145BB6960EF7333 CRC64;

Q9JVL8 Length: 173 May 30, 2002 09:32 Type: P Check: 1872 ..

1 MKDSLMDFFD FFAQGXKYKI ATEQGLDRKY FIFHYVCNL SAISSFCLMK
51 VIHSYTIICL LIMITIGAFG FILNRNVLIS VKNKQKFLSD MFPLEMPFFS
101 RYKIPILAR FLNRGMKNLY LQNFCLVMI CFSISILNE ERGMMPYFIC
151 YWGLFYFAN TYVFLSRDY KRI

11AA_SEQUENCE 1.0
ID Q9JUT37 PRELIMINARY; PRT; 172 AA.
AC Q9JUT37;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL INTEGRAL MEMBRANE PROTEIN.
GN NMA1989.

```

```
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagsels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajadream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RL meningitidis 22491."
EMBL: AL162757; CAB85209.1; .
KW Complete proteome.
SQ SEQUENCE 172 AA; 20428 MW; 0363D0A56B0692A1 CRC64;

09JUT37 Length: 172 May 30, 2002 09:32 Type: P Check: 492 ..

1 MKDSLMDAFD FFMQGNFKKI ATEQGLDRKY FIFQYIVCNL SAISSFCIMK
51 IHTYTLIYI LIAVIGIING FLIRSKVLST VKNKQKFLSD MFPLEMPFFF
101 VEYKIPILAR FLNRGMKMLY LPSEFCLVAI CFSIYILNEE GGMPMPYFACY
151 WEGLCYPVST YVFLSRDYK HI

11AA_SEQUENCE 1.0
ID P72121 PRELIMINARY; PRT; 723 AA.
AC P72121;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE OUTER MEMBRANE PROTEIN C PRECURSOR (OUTER MEMBRANE PROTEIN OPRC).
GN OPRC OR PA3790.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
EMBL: D28119; BAA05664.1; .
DR EMBL; AE004797; AAG07177.1; .
DR InterPro; IPR001005; Myb_DNA_bind.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00593; TonB_boxC; 1.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
KW Signal; Complete proteome.
FT SIGNAL 1 55 POTENTIAL.
FT CHAIN 56 723 MATURE PROTEIN C.
SQ SEQUENCE 723 AA; 79303 MW; E9FC33618E830A3E CRC64;
```

```
P72121 Length: 723 May 30, 2002 09:32 Type: P Check: 9154 ..

1 MEKRMSTQQR AAGNACPTAA FSPDPARLAQ RRRWAGAFAA LCGIALSPSA
51 LLAESHQHQ DHAVELAPSV VTGVAQSSPL TIVTNPKPR QPVASDGD
101 YLKTIPGFAY IRRGSGNGDP VLRGMFGSRL NILTNGGML GACPNRMDAP
151 TSYISPETYD KLVTKGPQT VLMGPGASAG TILFEREPER FGLGSRVNA
201 SLLAGSNGRF DAVLDAAGN RLGYLRTGN HAQSDDYEDG AGNTVPSRWK
251 KMNQDVAVGW TPDEDTLIEL TAGKGDGEAR YAGRGMDGSO FKRESLGLRF
301 VRSNVSDVLE KYEAQVYNNY ADHIMDNFRL RTPDPSSMMP MPMAQVDRR
351 TLGRLAATW RWDDEKLVGT VDAMRNEHRA RGSKYDMMTD YTTDADQFPW
401 SKDAVEHNYG AFGEIWFPAE ERDRLIGCLR LDRASVKDYR QTLKSGHMGH
451 AMANPTANDT RADTLPSEFV RYEHDLADSP TLLYAGLGHA ERFDPYWELE
501 SPKRGPNGSV NAFDKIKPEK TTQLDDEGLQY NGDKLQAWAS GYGVVQDFI
551 LFSYREGMG SSIQATNVDA RIMGELGAS YQLTGNWKT D ASLAYAWGKN
601 SSDDRALPQI PPLEARFGLT YEEGDSAGS LMRVAPQNR IARDQGNVYG
651 KDFDKSAGFG VESLNGAYRV TRNVKLSAGV DNLFDKDYTE HLNKAGDAGF
701 GFSANETVPE PORTFWTKVD FSF

11AA_SEQUENCE 1.0
ID Q916K4 PRELIMINARY; PRT; 392 AA.
AC Q916K4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE PROBABLE FATTY ACID DESATURASE.
GN PA0286.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AE004466; AAG03675.1; .
DR InterPro; IPR001522; Desaturase.
DR Pfam; PF01069; Desaturase; 1.
DR ProDom; PD002221; Desaturase; 1.
KW Complete proteome.
SQ SEQUENCE 392 AA; 44877 MW; 511E4B80E2F64928 CRC64;

Q916K4 Length: 392 May 30, 2002 09:32 Type: P Check: 1486 ..

1 MWENGFELDLS PMQLVAVTLV LTHITIVSVT VYLHRYSAHR ALDLHPALQH
51 FFRFWLWLT GNNTREMTAI HRKHAKCET VEDPHSPVHK GLFTVLRAGA
101 ELYKAEAKNQ DTLRIYGKNC PDDWIERNLV SRPIGGVVL MALIDLALFG
```

151 ALGLTWAVQ MWIPFWAG VVNGLHAVG YRNECRDAA TNLVPGILI
201 GGEELHNNH TYPNSAKLSV KKEFDLGWA WIKLEFSLRL ARVARVAPIA
251 HREVGKHSID MDTAMAILNN RFOIMAQYRK LVIAPLVKQE VAKADESVRH
301 LFERRAKRLLS RETSLQDRH HVRIDSMIAH SQALKVIYER RLALQIIMAK
351 TSANGHDMIA AIKDWVHEAE ASGIQSLKEF AAQLKTYSLR PA

11AA_SEQUENCE 1.0

ID Q9I5J3 PRELIMINARY; PRT; 358 AA.
AC Q9I5J3;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE HYPOTHETICAL PROTEIN PA0736.
GN PA0736.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Ewin A.L., Mizoguchi S.D., Warren P.,
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL: AE004508; AAG04125.1; ..
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 358 AA; 40019 MW; 035673ABE6EDDA58 CRC64;

O9I5J3 Length: 358 May 30, 2002 09:32 Type: P Check: 7512 ..

1 MQPIDRAQAQ QRANDILVFQ RELQRLDOEQ AFRLEPAQAR QLADYHLALL
51 DSYRDRFID HDLRSQLSL GMRVASFGA IAPFASVFL FYRFGLEPPT
101 VAQVAILLGS AFAAFATLM VQAKDASGYF SKLAAMVAF CTVLDTMLG
151 QIFNVTPSDL ALVPWALYAL LLAAYLCNARL LLAAILCYM GFIAARVGTW
201 GGGYWLVSVE RPENEFPAAA LIFAVPLCFD QNENSGFAVI YRVFALLGLF
251 LPMVLVLANWG SSGYLALPSA LIEGLYQVAG FVAALVIML GARRMADVS
301 NTGITEFVIF LYTKEFDWMW EAMPKYLFFL VLGSLALLIL LVLRLRLTPL
351 GIAARTDA

11AA_SEQUENCE 1.0

ID Q9CNB6 PRELIMINARY; PRT; 1230 AA.
AC Q9CNB6;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE RECB.
GN RECB OR PM0516.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PM70;
RX MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL: AE006087; AAK02600.1; ..
DR InterPro; IPR000212; UvrD-helicase.
DR Pfam; PF00580; UvrD-helicase; 3.
KW Complete proteome.
SQ SEQUENCE 1230 AA; 142199 MW; 4BEB673C75F76C71 CRC64;

O9CNB6 Length: 1230 May 30, 2002 09:32 Type: P Check: 8093 ..

1 MMDTAVLNVP TLPLNKVCLI EASAGTGKTY TIGSLYLRLL IQAGENSFSQ
51 PLTVEQIILVV TFTEATEEL KGRIRERIHQ AKKALIAQOE QGEQALQDDP
101 FLACIASIS DLDLAIQRLT IAEQTMDLAA IYTIHGFCRR MLMQYAFHSR
151 VHFNLTLNKD ETALLERLEK AFWREHFYSQ PFLVANYIHQ TLGSPQAVFS
201 ELROYIAQDL OPEPAYQAWL AMPLQDFLQO HIAPQOQNIQ QLKQOWLAOE
251 AEIQALILAE LEKTYPKGEK KRLKRTFEKK PNVLNWEKVI HEWATSPLVS
301 GLNDKLSKYF SOSALNTYAE EGATPDLSPV FALVEEVNAQ LDVQPFYAKL
351 LRYVYLRGVQ QTLIAVKAQH TEKNEDDLR LLRDALYSAQ GEELAQFIRV
401 QYPFAMIDEF QDTDAQYHI FSKYILHQOT TENGFIMIGD PKQAIYKFRG
451 ADIFTFYQAA EQADARFTLG TNWRSBQRLV NAVNSLFQFE QGLPEFLYPOI
501 QFLPVACQN KPTFNLNGQQ EPPFRGYVD VGVAKKNSGN LTSQKQTLA
551 TICARSIQW IQSAVGHDAI FYSAEAKQEE EKROPDLRAEK IAVLYKDWKE
601 ASFVSEALQK VGIASYVLSL KSNVFDGHEA QELALILTAC LHPRSERNIL
651 NAIATRFAL TTREISEIKQ DEQRTQVVE RFVNYQRIMQ WQGLVIMLHR
701 LFLDERIMEK LLSQVGERO TTDDLHLAEL LQEA STLNES AASLLRWFEX
751 QIQENRQEE QQIRLESERQ LVKIYTIHKS KGLEXYDLWL PFIDAPKPN
801 RALDPTYYHH EKQOVLMDLN ETHQTEIEOE QRAEAMRLFY VALTRAKYOV
851 VMALPETFVS HMNCLQYVLT QGAMTQTDVR AALTAFOQORV VYPPYKIQVE
901 EFELPIHLS TSIKENSADQ VLQCAEFHGN IERNQVTSF SAISALHEKT
951 MQLLTQAEER KTRDVSFLLD HKDYVSLGR NVALVPVAEM AGYSKGYTPE
1001 DEPAGTWGK VLRHYREKFP LDQPYDRAAV AQMCAQLQLE ETWLEPLQTW
1051 LTTILNTPLL RENHYTUSAL SAQDCIKEME FYLKFEHFEQ AHKFNQLQK
1101 YRFISAPLQI HTLKQGIKGL LRGFIDLVER YDGOYYLLDY KSNKLGTSPTS
1151 DYAPAHLOQV MLEQHYDMQY LFYTIALHRY LTLRDPHYQY ATHRGVLYT
1201 FLRGMGKQD QGISFHKPDA NLIQELBELF

11AA_SEQUENCE 1.0

ID Q9AAB3 PRELIMINARY; PRT; 208 AA.
AC Q9AAB3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN CC0691.
GN CC0691.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CBI5;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eissen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Uitterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of *Caulobacter crescentus*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005744; AAK22676.1; -.
 DR TIGR; CC0691; -.
 DR InterPro; IPR000182; Acetyltransf_GCN5.
 DR Pfam; PF00583; Acetyltransf_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 208 AA; 23051 MW; 41898AD0C249FC8E CRC64;

Q9AAB3 Length: 208 May 30, 2002 09:32 Type: P Check: 7467 ..

1 MPPIIVDNGP MLNLTADILQ GRYVRLPEVT VDHDDELKAA IDCDDPASWEI
 51 MSVNGCGEGF EDFWGLQGE TDRGERIGFA IRLVDGKVV GTSSYLNIIR
 101 LHGGLGIGAT FLNPEARSGP VNPESKRLML GHAFDKAGAI RVELVTDVRN
 151 ARSQAAIQKL GATKEGVLRN HKVTWGHVR DVAESITDY DMPAIRERLE
 201 FRLSETFA

11AA_SEQUENCE 1.0
 ID Q98JT9 PRELIMINARY; PRT; 153 AA.
 AC Q98JT9;
 DT 01-OCT-2001 (Tremblrel. 18, Created)
 DT 01-OCT-2001 (Tremblrel. 18, last sequence update)
 DT 01-OCT-2001 (Tremblrel. 18, last annotation update)
 DE MLI1786 PROTEIN.
 GN MLI1786.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP002998; BAB49076.1; -.
 KW Complete proteome.
 SQ SEQUENCE 153 AA; 17378 MW; 1C7F3793778B4964 CRC64;

Q98JT9 Length: 153 May 30, 2002 09:32 Type: P Check: 592 ..

1 MAIYRRKRD RHLQKAGEGI ACFFVGSANQ HAGMIAQRE GLRLCRGVL
 51 VFQRRLDLFA RGLVEGTAAG MRVPDQFTG VLKCHLTDV VRAMRLDPHR
 101 CRQHHIGHCS SRLWFSRYWR MLASRAASTI ARIAFSGVQ PSSASRQGRS
 151 AAW

11AA_SEQUENCE 1.0
 ID Q97KS8 PRELIMINARY; PRT; 260 AA.
 AC Q97KS8;

DT 01-OCT-2001 (Tremblrel. 18, Created)
 DT 01-OCT-2001 (Tremblrel. 18, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE ABC-TYPE SPERMIDINE/PUTRESCINE TRANSPORT SYSTEM, PERMEASE
 DE COMPONENT II.
 GN CAC0838.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium *Clostridium acetobutylicum*.";
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL; AE007599; AAK78814.1; -.
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp; 1.
 KW Complete proteome.
 SQ SEQUENCE 260 AA; 29101 MW; D9DC75197AFB77A0 CRC64;

Q97KS8 Length: 260 May 30, 2002 09:32 Type: P Check: 8409 ..

1 MVEKWLKRFY LTIYIFLYA PIVFLWFSE NSEKSSHWG HFSLTWYKAL
 51 LQDDRILFAL YTVLVVAIVS SIISTIFGTI SAIGISKMSR LPKLLNVN
 101 NIVYLNPDIV MAVSLMTLFI FLKIPFGLTT LIAHIAFSV PYVILSVLPK
 151 LTQLPTDIVK AALDLGATPS YAMRKIIIPQ IKSGIAGFL PAFMTSIDDF
 201 VISFNTGND VTNLSIEIYS MARRGITPEI NALSTLMPVT ILILLIANR
 251 KSIISKEKK

11AA_SEQUENCE 1.0
 ID Q97DB6 PRELIMINARY; PRT; 577 AA.
 AC Q97DB6;
 DT 01-OCT-2001 (Tremblrel. 18, Created)
 DT 01-OCT-2001 (Tremblrel. 18, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE MISMATCH REPAIR PROTEIN MUTS-LIKE ATPASE.
 GN CAC3563.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium *Clostridium acetobutylicum*.";
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL; AE007853; AAK81487.1; -.
 DR InterPro; IPR002106; AA_trna_ligase_II.
 DR InterPro; IPR000432; Muts_C.
 DR InterPro; IPR002863; Muts_N.
 DR Pfam; PF00488; Muts_C; 1.
 DR ProDom; PD001263; Muts_C; 1.
 DR SMART; SM00534; MUTSac; 1.
 DR SMART; SM00533; MUTSd; 1.
 DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.

KW Complete proteome.
SQ SEQUENCE 577 AA; 67131 MW; DDD636GDE85DEE7 CRC64;
Q97DB6 Length: 577 May 30, 2002 09:32 Type: P Check: 4438 ..
1 MDEDKLYMS AIALALGIS AFVEFQLAIG YKNYIYFLGT ALSIAAIFLA
51 FYMKAKIIHK LNYERFKRYW GEFQKRKRNT RNVKKFFEEH KEDNADEFNV
101 DEQTMEDLNM NKVELADRT LTPSGEEMLY KLEKTPFSK EKLERNSTII
151 KIFQNNKEVR EKVGLLEIKL GRKKENGVTD LIWKDIEVDY KYKYVENFLF
201 WGTLASILTI PILKGYIIL LAAFILMNTI EHNKFNKVE LYVQSLGYLN
251 GIINTANRIS KINCPEIKYV TDSLKATSSK LMKVAKKTAG IERVEGVDDI
301 GDALYNILPI EERKEFNSIN DIRKLRKELK EDYKALGEID ALMSIASFRQ
351 WIQYCEPEF VDNGRIRKCS EIYHPLLENP VNSIDLDDGK GIMLTGTNMA
401 GKSFTLRTIG LNSLLAQTIY TCAAKVYKTS FFKIMTSISP EDNISSGKSY
451 YFREAEALKR IINQCSDDRP VLCTIDEIFR GNPPIERVNA SAEILNYIEK
501 HMTLTLVATH DLELTEILKE DYLCYFSED IDDEGLQFDY KIKYGIICKTR
551 NAVKLKLYLE YPNEIIQKTN DRLAKIL
11AA_SEQUENCE 1.0
ID Q99VP6 PRELIMINARY; PRT; 501 AA.
AC Q99VP6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SA0682 PROTEIN (HYPOTHETICAL PROTEIN SAV0727).
GN SA0682 OR SAV0727.
OS Staphylococcus aureus (strain N315), and
OC Staphylococcus aureus (strain Mu50).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=158879, 158878;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain N315), and S.aureus (strain Mu50);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Muruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hiraoka H., Kihara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003131; BAB41915.1; -
DR EMBL; AP003360; BAB56889.1; -
DR InterPro; IPR000109; PTR2.
DR Pfam; PF00854; PTR2; 2.
DR PROSITE; PS01023; PTR2_2; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 501 AA; 54828 MW; 36F5C9DD0D3BD053 CRC64;
Q99VP6 Length: 501 May 30, 2002 09:32 Type: P Check: 6701 ..
1 MTQONSHGNQ IQDIPQTFE GHPRLGLVLF EYEFWERFSY YGNRALLIFY
51 MYRAVTDNGL GIDKTTAMSI MSVYGLIYM TSLPGGWIAAD RITGTRGATL
101 LGAVFIIGH ICISLPFALI GLFTSMFEII IGSGLMKPNI SNIVGRLYPE
151 NDRRDAGFV IFYMSVNMGA LSPILLOHF VNVKNFHGGF LIAAVGMALG

201 LWWYVLENRK NLGSGMKPT NPLTPAEKK YGLIGSVL AIYLIIVIGA
251 LTNSLSENLV SNTVLVLGIA LPIIYFTLII RSKDVTDER SRVKAEPILF
301 ILGMVFAAIQ EQGSNVLINIY GIEHSDMKLN LFGKTNFGE AIRQSINPLF
351 ILLAPIISL LMOKLGTQOP SLPVKFAIGT FLAGASYILI GIYVYASGSS
401 NFSVNWILS YIICVIGELC LSPGNSAAV KLAPEFNAQ MMSIWLITNA
451 SAQAINGLTV KLIPIQGTN YFIFGVVAI IYTIYLAFS PLIKAMKGI
501 R
11AA_SEQUENCE 1.0
ID Q92AV1 PRELIMINARY; PRT; 289 AA.
AC Q92AV1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LIN1818 PROTEIN.
GN LIN1818.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11267 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
DR EMBL; AL596170; CAC97049.1; -
DR Listlist; LIN01818; -
KW Complete proteome.
SQ SEQUENCE 289 AA; 32364 MW; C962CAA3D8FA8063 CRC64;
Q92AV1 Length: 289 May 30, 2002 09:32 Type: P Check: 1480 ..
1 MMKKVVKYIK HNGIVQGT VSARVGRNDV SGNAQLAYY MLFSIFPMLL
51 IAATLLAYLH IDKDSVFNNI KEFAPDQIMD FLEENLTLL TQKNGLLSI
101 GIATIMSAS NGMNAVYKSL NKAYGYTNKR NYVQRLLSM FFTLAMLATV
151 GATLLLVFG QQIGMELINH LNFESEDLSEF WNNLRWTVTL VVIFVETFL
201 YWVAPNRRST LISVLPGLAF STIGWTIASV GFAYYVNNFG NYSATYGSIG
251 VIIITMLMFY LTGIITMIGG ELNATIAIRK KKKELGEIN
11AA_SEQUENCE 1.0
ID Q929W8 PRELIMINARY; PRT; 391 AA.
AC Q929W8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LIN2155 PROTEIN.
GN LIN2155.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
BAguero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkut G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluteter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL: AL596171; CAC97385.1; -.
KW Complete proteome.
SQ SEQUENCE 391 AA; 43588 MW; F415BA5E7F99D506 CRC64;

Q929W8 Length: 391 May 30, 2002 09:32 Type: P Check: 3736 ..

1 MKATGIVEY NPFHNGHQLH LNKARELTKS DVIAVMGS FVORGEPAIT
51 PKWERTKMAL AAGVDMVIEL PVSFATQHAT IFAESIRLL DALHIDTLFF
101 GSEHGSEDF STAAKTIVEN EAAFNETIQL ALGDKTSTYA RAYTDAFTHL
151 FGKELDVTK PNNILGFHYA LAIQKQNPST ALQTMPREHS GYHDAEASHD
201 YIASATAIRK LLAGNLEEA SRYLPDSSIE VLNNYRGPF SLEDYWPULK
251 FRLIQASSDE LEGIRGVSEG IQNRMQLAAK KAHFSDFIE IMKTKRYSNA
301 RLQRTALQIL LNAQNTPPAE PYIRVLGMSK TGQKYLSLHK KNISLPIVTT
351 VSKAEPSSLK EDLRATDIYT LINGLEDYQA GDFHTPPILT L

11AA_SEQUENCE 1.0
ID Q92812 PRELIMINARY; PRT; 510 AA.
AC Q92812;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PGM PROTEIN.
GN PGM OR LIN2550.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
BAguero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkut G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluteter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL: AL596172; CAC97777.1; -.
KW Complete proteome.
SQ SEQUENCE 391 AA; 43588 MW; F415BA5E7F99D506 CRC64;

KW Complete proteome.
SQ SEQUENCE 510 AA; 56138 MW; E5A1A5A804B16A4D CRC64;

Q92812 Length: 510 May 30, 2002 09:32 Type: P Check: 2291 ..

1 MSKSPVAIIT LDGFGKRAET VGNAAVAQANK PNFDRYMANF PHGELKAAGL
51 DVGLEPGQMG NSEVGHNTNG AGRIVYQSLT RIDKALIEGE FOENKALNNA
101 FTHKENNSD LHLFGLSDG GYHSHINHLV ALLETAKDKG VKNVYIHAFL
151 DGRDVAPOSS LEYLETLEKA ISDLNYGAIA TVSGRFYAMD RDKRWERVEK
201 AYKAIVSAEG EKREDPIELV KASYANDKND EFVVPALITK DGKPVATVKD
251 NDVIFEFNER PDRAIQLSNA FTDKWDHFD RCANHPKNIK FVTMTLYNPS
301 IDAEVAFEP I EMKNVIGEVL SNEGSLQLRI AETEKYPHYT FPMNGGRNEE
351 FPGENRILIN SPKVETYDLK PEMSAYEVTD ALVEDIKNDK HDAILNFPAN
401 PDMVGHSGML EPTIKAIEAV DENLGRVVDL ILEKGSATI FADHGNSEFM
451 STEPGKPHTA HTTVPVPVIV TKKGVTLREG GRLADVAPTM LDLGVKKPPA
501 EMTGESLIQK

11AA_SEQUENCE 1.0
ID Q92937 PRELIMINARY; PRT; 315 AA.
AC Q92937;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-AUG-1998 (Tremblrel. 07, Last annotation update)
DE HYPOTHETICAL 36.7 KDA PROTEIN.
GN AF0925.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervilave A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL: AE001040; AAB90325.1; -.
DR TIGR: AF0925; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 315 AA; 36728 MW; 2855E42E710E7AAD CRC64;

Q92937 Length: 315 May 30, 2002 09:32 Type: P Check: 7306 ..

1 MLESWATSSL ILFVALVVI SLSLLYCSYL RRRRWKISG IFSVASFAVY
51 ILLEFWSFWN LLSNLLFILL VELAVFIILL PIFKVFEELK FENEKYGCEI
101 DGIPVITIGE KKKVYNNAVY TPLUKRIFFVT KSLKDVLSGE ELKAVIYHES
151 GHSKNKWMMI TRSTAMFWV LIAAVLTLTL FLLEMGKFOR NLKVSLEFILL
201 GALLIITYATF FWFVSWINEH EADLFAVKKS GYENFSKALF KTYFYNVLCGD

251 YAEFVGKIDL KNFNSGDVTP FEILKILKQ SIYLFPRNI LNQIPQTHP
301 PLRYRILLAH QTLKC
11AA_SEQUENCE 1.0
ID 029916 PRELIMINARY; PRT; 396 AA.
AC 029916;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE PHEROMONE SHUTDOWN PROTEIN (TRAB).
GN AF0331.
OS Archaeoglobus fulgidus.
OC Archaeae; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
Cotton M.D., Spriggs T., Arlanch P., Kaine B.P., Sykes S.M.,
Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL: AE001082; AAB90915.1; -
DR TIGR: AF0331; -
DR InterPro: IPR002816; TRAB.
DR Pfam: PF01963; TRAB; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 396 AA; 44188 MW; 859183D7432C8493 CRC64;
029916 Length: 396 May 30, 2002 09:32 Type: P Check: 4349 ..
1 MEERKLVIIVG TAHVSKRSVE EVAEVIEREK PDAVAVELCP RRYHALVHGQ
51 REETISADVT RKGNVMLLF QLILAYFQK VGEETGVKPG SEMLAIEKA
101 REAGADVLLI DRDIGLTFTFTR FMOKLTVEK IKLIHFLVRS TFGSDEIEVD
151 EMLEEDVLDL LKKEFRKISP NAAKVLIDER DYYMAANLN ALSRYNRIVA
201 VVGAGHRKGI EEALLLKEN PVNLRELEV KGRSYFKIF MGAFTALIVG
251 TPLIATSLN TEVLVOAFLY WFLINGVLSA VGASLARGHP LSIVTAFICA
301 WMTSLNPLVA AGWVSGLVEA WIRKPTVEDF SKLVEASLR EMMQNKFFRV
351 LVAALTNVG SMIGTIYGGW YIISTEGVDV AKVVGERVE VIGGLL
11AA_SEQUENCE 1.0
ID 09YFF0 PRELIMINARY; PRT; 110 AA.
AC 09YFF0;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE 110AA LONG HYPOTHETICAL BACTERIOFERRITIN COMIGRATORY PROTEIN.
GN APE0291.
OS Aeropyrum pernix.
OC Archaeae; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaei A., Kosugi H.,
Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000059; BAA79246.1; -
DR InterPro: IPR000866; Ahpc-TSA.
DR Pfam: PF00578; Ahpc-TSA; 1.
KW Complete proteome.
SQ SEQUENCE 110 AA; 12305 MW; E07B9D1AA31EE84D CRC64;
09YFF0 Length: 110 May 30, 2002 09:32 Type: P Check: 4613 ..
1 MLVSGDPAPD IEIQIDGST IRLSQLGRS VVLYFYPKAF TPQCTREAIQ
51 ENGLYEEFK LGAEVIGVSM DPPGRNRRFA QNYGVRFRLA SDVGEAFKS
101 FWGAQRPRTD
11AA_SEQUENCE 1.0
ID 09HIS0 PRELIMINARY; PRT; 150 AA.
AC 09HIS0;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
DE HYPOTHETICAL PROTEIN TAI1260.
GN TAI1260.
OS Thermoplasma acidophilum.
OC Archaeae; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;
OC Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumelster W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
acidophilum.";
RL Nature 407:508-513(2000).
DR EMBL: AL445067; CAC12384.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 150 AA; 17846 MW; A73F1605113594CC CRC64;
09HIS0 Length: 150 May 30, 2002 09:32 Type: P Check: 3314 ..
1 MFIITHLGF LLRYGFFLGF IGLRLLFPR FWLMFWLLDH YVLPCKGLY
51 LIEPYTARQI EMSYLELLRL DAKSLLSIC PCYCYPPLND EPGMINICY
101 LAFSLPGSA YDLDHVSHYN WERPVSVEFP QILAEHRFHL SVHMITLNYN

This Page Blank (uspto)